# GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2004, 15:56:54; Search time 9.48711 Seconds

(without alignments)

4199.299 Million cell updates/sec

Title: US-09-830-972-32

Perfect score: 705

Sequence: 1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVQKYSNSALG 141

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		8				
Result		Query				
No.	Score	Match	Length DB		ID	Description
1	680.5	96.5	522	3	AAY71312	Aay71312 Rat neuri
2	510	72.3	199	5	ABB81077	Abb81077 Rat neuro
3	503	71.3	118	5	ABB89192	Abb89192 Human pol
4	503	71.3	199	2	AAW53947	Aaw53947 Human NSP
5	503	71.3	199	2	AAW78313	Aaw78313 Fragment
6	503	71.3	199	2	AAY35903	Aay35903 Extended
7	503	71.3	199	3	AAB12805	Aab12805 Human NSP
8	503	71.3	199	4	AAB82348	Aab82348 Human NOG
9	503	71.3	199	5	ABG30939	Abq30939 Human Nog

		<b>5</b> 1 0	100	-	3 D D O 1 O O O	71-1-01000	***
10	503	71.3	199	5	ABB81080		Human neu
11	499.5	70.9	199	3	AAY71559		Rat Nogo
12	448	63.5	1178	3	AAY71311	<del>-</del>	Human neu
13	447	63.4	403	3	AAY71563		Rat Nogo
14	447	63.4	893	3	AAY95012		Human sec
15	447	63.4	983	6	ABU11573		Human MDD
16	447	63.4	1162	3	AAY71557	<del>-</del>	Rat Nogo
17	447	63.4	1163	3	AAY71310	<del>_</del>	Rat neuri
18	447	63.4	1163	3	AAY71384	Aay71384	Alternati
19	447	63.4	1163	5	ABB81074	Abb81074	Rat neuro
20	447	63.4	1192	3	AAY56967	Aay56967	Human MAG
21	447	63.4	1192	4	AAB82349	Aab82349	Human NOG
22	447	63.4	1192	4	AAU04591	Aau04591	Human Nog
23	447	63.4	1192	5	ABG30938	Abg30938	Human Nog
24	447	63.4	1192	5	ABP68600	Abp68600	Human pan
25	447	63.4	1192	5	ABB81078	Abb81078	Human neu
26	447	63.4	1192	6	ABR59667	Abr59667	Human Nog
27	443	62.8	103	4	AAE03980	Aae03980	Human gen
28	443	62.8	200	4	AAB64514	Aab64514	Human sec
29	443	62.8	359	3	AAY71558	Aay71558	Rat Nogo
30	443	62.8	360	3	AAY71383	Aay71383	Rat neuri
31	443	62.8	360	4	AAE03987	Aae03987	Human gen
32	443	62.8	360	5	ABB81076	Abb81076	Rat neuro
33	443	62.8	361	3	AAY71385	Aay71385	Alternati
34	443	62.8	373	.3	AAY53624		A bone ma
35	443	62.8	373	3	AAY56969	Aay56969	Human MAG
36	443	62.8	373	3	AAB24242	Aab24242	Human Nog
37	443	62.8	373	4	AAB82350		Human NOG
38	443	62.8	373	5	AAM47954	Aam47954	Human RTN
39	443	62.8	373	5	ABG30937	Abq30937	Human Nog
40	443	62.8	373	5	ABP68601	Abp68601	Human pan
41	443	62.8	373	5	ABB81079		Human neu
42	443	62.8	379	7	ADB85283	Adb85283	Rat fooce
43	440	62.4	91	2	AAY12360	Aav12360	Human 5'
44	439	62.3	291	4	AAM93484	<del>-</del>	Human pol
45	410.5	58.2	642	2	AAW58383		Human sec
. ~				_			

## ALIGNMENTS

```
RESULT 1
AAY71312
    AAY71312 standard; protein; 522 AA.
ID
XX
AC
    AAY71312;
XX
DT
     02-NOV-2000 (first entry)
XX
DE
    Rat neurite growth inhibitor Nogo C.
XX
     Rat; neurite growth inhibitor; Nogo C; neural cell; myelin; CNS;
KW
     central nervous system; neoplastic disease; antiproliferative; glioma;
KW
     antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;
KW
     degenerative nerve disease; Alzheimer's disease; Parkinson's disease;
KW
     hyperproliferative disorder; benign dysproliferative disorder; diagnosis;
ΚW
     psoriasis; tissue hypertrophy; neuronal regeneration; treatment;
KW
```

```
KW
     structural plasticity; screening.
XX
OS
     Rattus sp.
XX
                      Location/Qualifiers
FH
     Key
                      1. .39
FT
     Region
                      /note= "Sequence upstream to the N-terminus of Nogo C
FT
FT
                      protein"
FT
     Misc-difference 3
FT
                      /note= "Encoded by TAG"
FT
     Region
                      11. .191
                      /note= "Region specifically described in claim 16"
FT
FT
     Misc-difference 29
                      /note= "Encoded by TAA"
FT
FT
     Protein
                      40. .238
                      /label= Nogo_C_protein
FT
FT
     Region
                      51. .238
                      /note= "C-terminal common region found in Nogo A, B and C
FT
FT
                      isoforms "
FT
     Region
                      239. .522
                      /note= "Sequence downstream to the C-terminus of Nogo C
FT
FT
                      protein"
     Misc-difference 239
FT
FT
                      /note= "Encoded by TGA"
FT
     Misc-difference 263
                      /note= "Encoded by TGA"
FT
FT
     Misc-difference 276
FT
                      /note= "Encoded by TAG"
     Misc-difference 281
FT
                      /note= "Encoded by TGA"
FT
FT
     Misc-difference 295
FT
                      /note= "Encoded by TAA"
     Misc-difference 298
FT
FT
                      /note= "Encoded by TAA"
FT
     Misc-difference 314
                      /note= "Encoded by TGA"
FT
FT
     Misc-difference 318
                      /note= "Encoded by TGA"
FT
FT
     Misc-difference 335
FT
                      /note= "Encoded by TAG"
FT
     Misc-difference 371
                      /note= "Encoded by TGA"
FT
FT
     Misc-difference 374
                      /note= "Encoded by TAG"
FT
FT
     Misc-difference 380
FT
                      /note= "Encoded by TAA"
FT
     Misc-difference 406
FT
                      /note= "Encoded by TAA"
FT
     Misc-difference 408
FT
                      /note= "Encoded by TAG"
     Misc-difference 410
FT
                      /note= "Encoded by TAA"
FT
     Misc-difference 422
FT
                      /note= "Encoded by TAA"
FT
     Misc-difference 433
FT
FT
                      /note= "Encoded by TGA"
```

FT

Misc-difference 440

```
FT
                     /note= "Encoded by TAG"
FT
     Misc-difference 453
FT
                     /note= "Encoded by TAG"
FT
     Misc-difference 465
FT
                     /note= "Encoded by TAG"
FT
     Misc-difference 482
FТ
                     /note= "Encoded by TAG"
FT
     Misc-difference 513
FT
                     /note= "Encoded by TAA"
XX
PN
     W0200031235-A2.
XX
PD
     02-JUN-2000.
XX
PF
     05-NOV-1999;
                    99WO-US026160.
XX
PR
     06-NOV-1998;
                    98US-0107446P.
XX
PA
     (SCHW/) SCHWAB M E.
PA
     (CHEN/) CHEN M S.
XX
PΙ
     Schwab ME,
                 Chen MS;
XX
DR
     WPI; 2000-400052/34.
DR
     N-PSDB; AAD01175.
XX
PT
     Nogo proteins and nucleic acids useful for treating neoplastic disorders
PT
     of the central nervous system and inducing regeneration of neurons.
XX
PS
     Claim 7; Fig 14; 122pp; English.
XX
CC
     The present sequence is a rat Nogo C protein which is a potent neural
CC
     cell growth inhibitor and is free of all central nervous system (CNS)
CC
     myelin material with which it is natively associated. Nogo proteins and
CC
     fragments displaying neurite growth inhibitory activity are used in the
     treatment of neoplastic disease of the CNS e.g. glioma, glioblastoma,
CC
CC
     medulloblastoma, craniopharyngioma, ependyoma, pinealoma,
CC
     haemangioblastoma, acoustic neuroma, oligodendroglioma, menagioma,
CC
     neuroblastoma or retinoblastoma and degenerative nerve diseases e.g.
CC
     Alzheimer's and Parkinson's diseases. Therapeutics which promote Nogo
CC
     activity can be used to treat or prevent hyperproliferative or benign
CC
     dysproliferative disorders e.g. psoriasis and tissue hypertrophy.
CC
     Ribozymes or antisense Nogo nucleic acids can be used to inhibit
CC
     production of Nogo protein to induce regeneration of neurons or to
CC
     promote structural plasticity of the CNS in disorders where neurite
CC
     growth, regeneration or maintenance are deficient or desired. The animal
     models can be used in diagnostic and screening methods for predisposition
CC
CC
     to disorders and to screen for or test molecules which can treat or
CC
     prevent disorders or diseases of the CNS. Note: SEQ ID numbers 35-42 are
CC
     referred in claim 32 and SEQ ID NO: 29 in disclosure of the
CC
     specification. However the specification does not include sequences for
CC
     these SEQ ID numbers
XX
SQ
     Sequence 522 AA;
  Query Match
                          96.5%; Score 680.5; DB 3; Length 522;
```

Best Local Similarity 98.6%; Pred. No. 2.5e-71;

```
Matches 138; Conservative
                               0;
                                  Mismatches
                                                   Indels
                                                1;
                                                             1:
                                                                        1;
                                                                 Gaps
           3 SGEAGVSCLRENFAVYSVSVGMHNL-LLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKT 61
Qу
             Db
           4 SGEAGVSCLRENFAVYSVSVGMHNLXLLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKT 63
Qу
          62 GVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSDEGHPFRAYL 121
             Db
          64 GVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYL 123
         122 ESEVAISEELVOKYSNSALG 141
Qν
             Db
         124 ESEVAISEELVQKYSNSALG 143
RESULT 2
ABB81077
    ABB81077 standard; protein; 199 AA.
XX
AC
    ABB81077;
XX
DT
    05-NOV-2002 (first entry)
XX
DE
    Rat neurotransmitter receptor protein Nogo-C.
XX
KW
    Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;
KW
    central nervous system; peripheral nervous system; tranquillizer; Nogo;
KW
    vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;
KW
    nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;
KW
    osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;
KW
    neurotransmitter receptor; rat; receptor.
XX
OS
    Rattus norvegicus.
XX
PN
    US2002072493-A1.
XX
    13-JUN-2002.
PD
XX
PF
    28-JUN-2001; 2001US-00893348.
XX
PR
    19-MAY-1998;
                  98IL-00124500.
PR
                  98WO-US014715.
    21-JUL-1998;
                  98US-00218277.
PR
    22-DEC-1998;
                  99US-00314161.
PR
    19-MAY-1999;
XX
PA
    (YEDA ) YEDA RES & DEV CO LTD.
XX
PΙ
    Eisenbach-Schwartz M.
                         Hauben E, Cohen IR,
                                             Beserman P, Mosonego A;
    Moalem G:
PΙ
XX
DR
    WPI; 2002-607255/65.
DR
    N-PSDB; ABN86600.
XX
PΤ
    Promoting nerve regeneration and preventing neuronal degeneration in the
    central/peripheral nervous system from injury/disease, comprises
PT
PT
    administering nervous system-specific activated T cells/antigen, or
PΤ
    analogs/peptides.
```

```
PS
     Example 5; Page 48-49; 93pp; English.
XX
CC
     The invention relates to promoting nerve regeneration or conferring
CC
     neuroprotection and preventing or inhibiting neuronal degeneration in the
CC
     central/peripheral nervous system (NS). The method involves administering
CC
     NS-specific activated T cells, NS-specific antigen, its analogue or its
CC
     peptide, a nucleotide sequence the NS-specific antigen or its analogue or
CC
     combinations. The method is useful for promoting nerve regeneration and
CC
     preventing neuronal degeneration in central/peripheral nervous system
CC
     from injury/disease, where the injury is spinal cord injury, blunt
CC
     trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or
CC
     damages caused by surgery such as tumour excision. The disease is not an
CC
     autoimmune disease or neoplasm. The disease results in a degenerative
     process occurring in either gray or white matter or both. The disease is
CC
CC
     diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's
     disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,
CC
CC
     amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and
CC
     vitamin deficiency, intervertebral disc herniation, prion diseases such
     as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral
CC
CC
     neuropathies associated with various diseases, including but not limited
CC
     to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute
CC
     sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary
CC
     amyloidosis, obstructive lung diseases, acromegaly, malabsorption
CC
     syndromes, polycythemia vera, immunoglobulin (Ig) A- and IgG gamma-
     pathies, complications of various drugs (e.g., metronidazole) and toxins
CC
CC
     (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia
CC
     telangectasia, Friedreich's ataxia, amyloid polyneuropathies,
CC
     adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's
CC
     disease, or lipoproteinemia. The present sequence represents the rat
CC
     neurotransmitter receptor protein Nogo-C, an example of NS-specific
CC
     antigen
XX
     Sequence 199 AA;
SO
 Query Match
                         72.3%;
                                 Score 510; DB 5;
                                                   Length 199;
 Best Local Similarity
                         99.0%;
                                 Pred. No. 9.6e-52;
 Matches 103; Conservative
                                0; Mismatches
                                                 1; Indels
                                                               0; Gaps
                                                                           0;
Qу
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
             Db
           1 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qy
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 3
ABB89192
    ABB89192 standard; protein; 118 AA.
XX
AC
    ABB89192;
XX
DT
    24-MAY-2002 (first entry)
XX
DE
    Human polypeptide SEQ ID NO 1568.
```

XX

```
XX
KW
     Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW
     antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
     vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW
KW
     cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
     neurological disease; infection; human; secreted protein.
KW
XX
OS
     Homo sapiens.
XX
PN
     W0200190304-A2.
XX
     29-NOV-2001.
PD
XX
ΡF
     18-MAY-2001; 2001WO-US016450.
XX
PR
     19-MAY-2000; 2000US-0205515P.
XX
PΑ
     (HUMA-) HUMAN GENOME SCI INC.
XX
PΙ
     Birse CE, Rosen CA;
XX
DR
     WPI; 2002-122018/16.
DR
    N-PSDB; ABL89601.
XX
PT
     Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT
     prevention of neural, immune system, muscular, reproductive,
     gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT
PT
     disorders.
XX
PS
     Claim 11; SEQ ID NO 1568; 2081pp + Sequence Listing; English.
XX
CC
    The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC
     (ABB89040-ABB90444) useful for preventing, treating or ameliorating
    medical conditions e.g. by protein or gene therapy. The genes are
CC
CC
     isolated from a range of human tissues disclosed in the specification.
CC
    The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC
     the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC
    ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC
    breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC
    disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC
     anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC
    multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC
     cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC
     ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CÇ
    infectious diseases such as viral, bacterial, fungal and parasitic
     infections. Note: The sequence data for this patent did not form part of
CC
CC
     the printed specification, but was obtained in electronic format directly
CC
     from WIPO at ftp.wipo.int/pub/published pct sequences
XX
SO
    Sequence 118 AA;
                                  Score 503; DB 5; Length 118;
  Query Match
                          71.3%;
  Best Local Similarity
                          98.1%; Pred. No. 3.2e-51;
 Matches 102; Conservative
                                 1; Mismatches
                                                   1; Indels
                                                                 0; Gaps
```

```
Db
            1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
           98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qy
              Db
           61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 4
AAW53947
     AAW53947 standard; protein; 199 AA.
XX
AC
    AAW53947;
XX
DТ
    24-JUL-1998 (first entry)
XX
DΕ
    Human NSPLP protein A.
XX
KW
    NSPLP; neuroendocrine-specific protein-like protein; human; gene therapy;
KW
    neurodegenerative disease; amyotrophic lateral sclerosis; cancer.
XX
OS
    Homo sapiens.
XX
PN
    WO9806841-A2.
XX
PD
    19-FEB-1998.
XX
PF
    24-JUL-1997;
                   97WO-US013469.
XX
PR
    12-AUG-1996;
                   96US-00700607.
XX
PΑ
     (INCY-) INCYTE PHARM INC.
XX
PΙ
    Bandman O, Au-Young J, Goli SK, Hillman J;
XX
DR
    WPI; 1998-159533/14.
DR
    N-PSDB; AAV23695.
XX
PT
    Human neuro-endocrine-specific protein-like proteins - useful for
PT
    diagnosis, monitoring and treatment of cancer and neuro-degenerative
PΤ
    disease.
XX
PS
    Claim 1; Page 38; 73pp; English.
XX
CC
    This sequence is a human neuroendocrine-specific protein-like protein
CC
    (NSPLP) of the invention. Recombinant cells transformed with the DNA are
CC
    used to express the NSPLP proteins, which are used to treat cancer and
CC
    neurodegenerative diseases such as amyotrophic lateral sclerosis. Also
CC
    antisense nucleic acids and antagonists of NSPLP can be used to inhibit
CC
    activity of the NSPLP proteins. Antibodies specific for NSPLP are used
CC
    for diagnosis and monitoring treatment of diseases associated with NSPLP
CC
    expression, in usual immunoassays, and to isolate NSPLP from natural
CC
    sources. The NSPLP proteins, or their fragments can also be used in drug
CC
    screening to identify NSPLP antagonists. The nucleic acid can be used
CC
    diagnostically and for monitoring treatment (in hybridisation or
CC
    amplification assays); to isolate closely related sequences; in gene
CC
    therapy for both sense and antisense applications (including use of
```

ribozymes) and for mapping the natural genomic sequence

CC

```
XX
SQ
     Sequence 199 AA;
  Query Match
                         71.3%; Score 503; DB 2; Length 199;
  Best Local Similarity
                         98.1%;
                                 Pred. No. 6.4e-51;
  Matches 102; Conservative
                                1; Mismatches
                                                 1;
                                                    Indels
                                                               0;
                                                                           0;
                                                                   Gaps
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
              Db
           1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVOKYSNSALG 141
Qy
              Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 5
AAW78313
ID
    AAW78313 standard; protein; 199 AA.
XX
AC
    AAW78313;
XX
DT
    13-APR-1999
                 (first entry)
XX
DE
    Fragment of human secreted protein encoded by gene 69.
XX
KW
    Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW
    diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW
    developmental abnormality; foetal deficiency; blood; allergy; renal;
KW
    immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW
    inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
ΚW
    cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW
    osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW
    endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS
    Homo sapiens.
XX
PN
    WO9856804-A1.
XX
PD
    17-DEC-1998.
XX
                   98WO-US012125.
PF
    11-JUN-1998;
XX
                   97US-0049547P.
PR
    13-JUN-1997;
    13-JUN-1997;
                   97US-0049548P.
PR
PR
    13-JUN-1997;
                   97US-0049549P.
PR
                   97US-0049550P.
    13-JUN-1997;
PR
    13-JUN-1997;
                   97US-0049566P.
PR
    13-JUN-1997;
                   97US-0049606P.
PR
                   97US-0049607P.
    13-JUN-1997;
PR
    13-JUN-1997;
                   97US-0049608P.
                   97US-0049609P.
PR
    13-JUN-1997;
                   97US-0049610P.
PR
    13-JUN-1997;
PR
                   97US-0049611P.
    13-JUN-1997;
PR
    13-JUN-1997;
                   97US-0050901P.
PR
    13-JUN-1997;
                   97US-0052989P.
PR
    08-JUL-1997;
                   97US-0051919P.
```

```
PR
    18-AUG-1997;
                   97US-0055984P.
PR
    12-SEP-1997;
                   97US-0058665P.
PR
    12-SEP-1997;
                   97US-0058668P.
PR
    12-SEP-1997;
                   97US-0058669P.
PR
    12-SEP-1997;
                   97US-0058750P.
PR
    12-SEP-1997;
                   97US-0058971P.
PR
                   97US-0058972P.
    12-SEP-1997;
                   97US-0058975P.
PR
    12-SEP-1997;
PR
    02-OCT-1997;
                   97US-0060834P.
PR
    02-OCT-1997;
                   97US-0060841P.
                   97US-0060844P.
PR
    02-OCT-1997;
                   97US-0060865P.
PR
    02-OCT-1997;
PR
                   97US-0061059P.
    02-OCT-1997;
PR
    02-OCT-1997;
                   97US-0061060P.
XX
PΑ
     (HUMA-) HUMAN GENOME SCI INC.
XX
ΡI
    Moore PA, Shi Y, Rosen CA, Ruben SM, Lafleur DW, Olsen HS;
ΡI
    Ebner R, Brewer LA, Young P, Greene JM, Ferrie AM, Yu G, Ni J;
PΙ
    Feng P;
XX
DR
    WPI; 1999-080881/07.
    N-PSDB; AAX04379.
DR
XX
PT
    New isolated human genes and the secreted polypeptides they encode -
PT
    useful for diagnosis and treatment of e.g. cancers, neurological
PT
    disorders, immune diseases, inflammation or blood disorders.
XX
PS
    Disclosure; Page 62; 380pp; English.
XX
CC
    This sequence represents a fragment of a secreted human protein encoded
CC
    by the nucleic acid molecule detailed in the descriptor line. The gene
CC
    can be used to generate fusion proteins by linking to the gene to a human
CC
    immunoglobulin Fc portion (e.g. AAX04302) for increasing the stability of
CC
    the fused protein as compared to the human protein only. The invention
CC
    relates to 86 novel genes and their fragments (nucleic acid sequences:
    AAX04311-X04410; amino acid sequences AAW78126-W78225) which are useful
CC
CC
    for preventing, treating or ameliorating medical conditions e.g. by
CC
    protein or gene therapy. Also, pathological conditions can be diagnosed
CC
    by determining the amount of the new polypeptides in a sample or by
    determining the presence of mutations in the new polynucleotides.
CC
    Specific uses are described for each of the 86 polynucleotides, based on
CC
CC
    which tissues they are most highly expressed in (see AAX04311 for
CC
    described uses)
XX
SO
    Sequence 199 AA;
 Query Match
                         71.3%; Score 503; DB 2; Length 199;
 Best Local Similarity
                         98.1%;
                                Pred. No. 6.4e-51;
 Matches 102; Conservative
                               1; Mismatches
                                                 1; Indels
Qу
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
             Db
           1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
```

XX

```
RESULT 6
AAY35903
ID
     AAY35903 standard; protein; 199 AA.
XX
AC
     AAY35903;
XX
DT
     13-SEP-1999 (first entry)
XX
DE
     Extended human secreted protein sequence, SEQ ID NO. 152.
XX
KW
     Secreted protein; human; cytokine; cellular proliferation; cell movement;
KW
     cellular differentiation; immune system regulator; anti-inflammatory;
     haematopoiesis regulator; tissue growth regulator; tumour inhibitor;
KW
KW
     reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;
KW
     genetic disease.
XX
OS
     Homo sapiens.
XX
ÞΝ
     W09931236-A2.
XX
PD
     24-JUN-1999.
XX
PF
     17-DEC-1998;
                    98WO-IB002122.
XX
PR
                    97US-0069957P.
     17-DEC-1997;
PR
     09-FEB-1998;
                    98US-0074121P.
PR
     13-APR-1998;
                    98US-0081563P.
PR
     10-AUG-1998;
                    98US-0096116P.
XX
PA
     (GEST ) GENSET.
XX
PΤ
     Bougueleret L, Duclert A, Dumas Milne Edwards J;
XX
DR
     WPI; 1999-385906/32.
     N-PSDB; AAX97587.
DR
XX
PT
     New isolated human secreted proteins.
XX
PS
     Claim 9; Page 185-186; 516pp; English.
XX
CC
     This sequence is encoded by an extended human secreted protein coding
CC
     sequence of the invention. The secreted proteins can be used in treating
CC
     or controlling a variety of human conditions. The secreted proteins may
CC
     act as cytokines or may affect cellular proliferation or differentiation
CC
     or may act as immune system regulators, haematopoiesis regulators, tissue
CC
     growth regulators, regulators of reproductive hormones or cell movement
CC
     or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or
     tumour inhibition activity. The DNAs can be used in forensic procedures
CC
CC
     to identify individuals or in diagnostic procedures to identify
CC
     individuals having genetic diseases resulting from abnormal expression of
CC
     the genes corresponding to the extended cDNAs. They are also useful for
CC
     constructing a high resolution map of the human chromosomes. They can
CC
     also be used for gene therapy to control or treat genetic diseases
```

```
SO
     Sequence 199 AA;
  Query Match
                        71.3%; Score 503; DB 2; Length 199;
  Best Local Similarity
                        98.1%; Pred. No. 6.4e-51;
  Matches 102; Conservative
                              1; Mismatches
                                              1; Indels
                                                             0; Gaps
                                                                        0;
Qу
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
             1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Db
Qу
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 7
AAB12805
    AAB12805 standard; protein; 199 AA.
XX
AC
    AAB12805;
XX
DТ
    24-NOV-2000 (first entry)
XX
DE
    Human NSPH protein sequence SEQ ID NO:4.
XX
KW
    Human; neuroendocrine-specific protein; NSPH; NSPA; NSPB; NSPC.
XX
OS
    Homo sapiens.
XX
PN
    CN1253180-A.
XX
PD
    17-MAY-2000.
XX
PF
    30-OCT-1998;
                  98CN-00121473.
XX
PR
    30-OCT-1998;
                  98CN-00121473.
XX
PA
    (UYFU-) UNIV FUDAN.
XX
PΙ
    Yu L, Zhao Y, Zhang H;
XX
DR
    WPI; 2000-466537/41.
    N-PSDB; AAA72981.
DR
XX
PT
    Specific protein of human neuroendocrine, coding sequence and its
PT
    preparating process and application.
XX
PS
    Claim 4; Page 14-15; 21pp; Chinese.
XX
CC
    The present invention relates to a new member of the human neuroendocrine
CC
    specific protein family, designated NSPH. The present sequence represents
CC
    the human NSPH protein
XX
SO
    Sequence 199 AA;
 Query Match
                        71.3%; Score 503; DB 3; Length 199;
```

98.1%; Pred. No. 6.4e-51;

Best Local Similarity

```
Matches 102; Conservative 1; Mismatches
                                                 1; Indels
                                                               0;
                                                                   Gaps
                                                                           0;
           38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
              1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Db
           98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
              61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
Db
RESULT 8
AAB82348
ΙD
     AAB82348 standard; protein; 199 AA.
XX
AC
     AAB82348;
XX
DT
     23-JUL-2001 (first entry)
XX
DE
     Human NOGO-C protein.
XX
     NOGO-C; human; chromosome 2p21; neuropathy; spinal injury; brain injury;
KW
     stroke; neuronal degeneration; Alzheimer's disease; Parkinson's disease;
KW
     neuromuscular disorder; psychiatric disorder; developmental disorder;
KW
KW
     neuroprotective; nootropic; neuroleptic; antiparkinsonian;
KW
     cerebroprotective; neuroleptic; diagnosis; therapy.
XX
OS
     Homo sapiens.
XX
PN
     W0200136631-A1.
XX
PD
     25-MAY-2001.
XX
PF
     14-NOV-2000; 2000WO-GB004345.
XX
PR
     15-NOV-1999;
                   99GB-00026995.
PR
     24-JAN-2000; 2000GB-00001550.
XX
PΑ
     (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI
    Michalovich D, Prinjha R;
XX
DR
    WPI; 2001-343822/36.
DR
    N-PSDB; AAF90323.
XX
PΤ
    New polypeptide designated NOGO-C is a splice variant of the human NOGO
PT
    gene and may be useful in the treatment of neural disorders including
PT
    Alzheimer's and Parkinson's diseases.
XX
PS
    Claim 3; Page 25; 25pp; English.
XX
CC
    The present sequence is that of human NOGO-C, encoded by a novel splice
CC
    variant of the human NOGO gene on chromosome 2p21. 2 Other splice
    variants, NOGO-A and NOGO-B, have previously been identified. The
CC
    invention provides NOGO-C polypeptides and polynucleotides, and methods
CC
CC
    for producing such polypeptides by recombinant techniques. Also disclosed
CC
    are methods for utilising NOGO-C polypeptides and polynucleotides in the
```

```
CC
     treatment of diseases including neuropathies, spinal injury, brain
CC
     injury, stroke, neuronal degeneration, for example Alzheimer's disease
     and Parkinson's disease, neuromuscular disorders, psychiatric disorders
CC
CC
     and developmental disorders. Also provided are methods for identifying
CC
     agonists and agonists for use in treating conditions associated with NOGO
     -C imbalance, and diagnostic assays for detecting diseases associated
CC
CC
     with inappropriate NOGO-C activity or levels
XX
SQ
     Sequence 199 AA;
  Query Match
                         71.3%; Score 503; DB 4; Length 199;
                         98.1%; Pred. No. 6.4e-51;
  Best Local Similarity
  Matches 102; Conservative
                               1; Mismatches
                                                1; Indels
                                                               0; Gaps
                                                                          0;
           38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
              Db
            1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
           98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qy
              Db
           61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 9
ABG30939
     ABG30939 standard; protein; 199 AA.
ID
XX
AC
     ABG30939;
XX
DΤ
     21-OCT-2002 (first entry)
XX
DΕ
     Human NogoC protein.
XX
KW
     Human; Nogo; BACE; acute neuronal injury; spinal injury; head injury;
KW
     stroke; peripheral nerve damage; neoplastic disorder; glioblastoma;
KW
     neuroblastoma; hyperproliferative disorder; dysproliferative disorder;
KW
     cirrhosis; psoriasis; keloid formation; fibrocystic condition; cancer;
     tissue hypertrophy; central nervous system; axon regeneration; NogoC;
KW
     Nogo-associated disease; metastasis.
KW
XX
OS
    Homo sapiens.
XX
ΡN
    WO200257483-A2.
XX
PD
     25-JUL-2002.
XX
PF
    18-JAN-2002; 2002WO-GB000228.
XX
PR
    18-JAN-2001; 2001GB-00001312.
XX
PA
     (GLAX ) GLAXO GROUP LTD.
     (SMIK ) SMITHKLINE BEECHAM PLC.
PA
XX
PΙ
    Blackstock WP, Hale RS, Prinjha R, Rowley A;
XX
DR
    WPI; 2002-599722/64.
DR
    N-PSDB; ABK90135.
```

```
XX
PT
     Identifying modulators of Nogo or BACE activity for treating acute
PT
     neuronal injuries, neoplastic or dysproliferative disorders, comprises
PT
     providing and monitoring interaction between Nogo and BACE polypeptides.
XX
PS
     Disclosure; Page 64; 68pp; English.
XX
CC
     The present invention relates to a new method of identifying modulators
CC
     of Nogo function or BACE activity. The method involves providing Nogo and
CC
     BACE polypeptides capable of binding with each other, monitoring the
CC
     interaction between these polypeptides, and determining if the test agent
CC
     is a modulator of Nogo or BACE activity. The method is useful in treating
CC
     acute neuronal injuries, such as spinal or head injury, stroke,
CC
     peripheral nerve damage, and in neoplastic (e.g. glioblastomas,
CC
     neuroblastomas), hyperproliferative or dysproliferative disorders (e.g.
CC
     cirrhosis, psoriasis, keloid formation, fibrocystic conditions, tissue
CC
     hypertrophy) of the central nervous system. The BACE polypeptide is
     useful in screening methods to identify agents that may act as modulators
CC
CC
     of BACE activity and in particular agents that may be useful in treating
     Nogo-associated diseases. The modulators of Nogo or BACE polypeptides,
CC
     and the polynucleotide encoding the BACE polypeptide are useful in
CC
     manufacturing a medicament for the treatment or prevention of disorders
CC
CC
     responsive to the modulation of Nogo activity, in alleviating the
     symptoms or improving the condition of a patient suffering from this
CC
CC
     disorder, in axon regeneration, or in preventing metastasis or spreading
CC
     of a cancer. The polynucleotide may also be an essential component in
CC
     assays, a probe, in recombinant protein synthesis, and in gene therapy
CC
     techniques. The present amino acid sequence represents the human NogoC
CC
     protein of the invention
XX
SQ
     Sequence 199 AA;
  Query Match
                         71.3%;
                                 Score 503; DB 5; Length 199;
  Best Local Similarity
                         98.1%;
                                 Pred. No. 6.4e-51;
  Matches 102; Conservative
                                1; Mismatches
                                                 1; Indels
                                                               0;
                                                                   Gaps
                                                                           0;
Qу
           38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
              Db
            1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Qy
           98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
              Db
           61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 10
ABB81080
    ABB81080 standard; protein; 199 AA.
XX
AC
    ABB81080;
XX
DT
    05-NOV-2002 (first entry)
XX
DE
    Human neurotransmitter receptor protein Nogo-C.
XX
    Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;
KW
     central nervous system; peripheral nervous system; tranquillizer; Nogo;
KW
```

KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant; nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic; KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy; KW KW neurotransmitter receptor; human; receptor. XX OS Homo sapiens. XX PN US2002072493-A1. XX PD13-JUN-2002. XX PF 28-JUN-2001; 2001US-00893348. XX PR 19-MAY-1998; 98IL-00124500. PR 21-JUL-1998; 98WO-US014715. PR 22-DEC-1998; 98US-00218277. PR 19-MAY-1999; 99US-00314161. XX PA (YEDA ) YEDA RES & DEV CO LTD. XX Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonego A; PΙ PΙ Moalem G; XX DR WPI; 2002-607255/65. DR N-PSDB; ABN86601. XX PT Promoting nerve regeneration and preventing neuronal degeneration in the central/peripheral nervous system from injury/disease, comprises PTPTadministering nervous system-specific activated T cells/antigen, or PTanalogs/peptides. XX PS Example; Page 57-58; 93pp; English. XX CC The invention relates to promoting nerve regeneration or conferring CC neuroprotection and preventing or inhibiting neuronal degeneration in the CC

central/peripheral nervous system (NS). The method involves administering NS-specific activated T cells, NS-specific antigen, its analogue or its peptide, a nucleotide sequence the NS-specific antigen or its analogue or combinations. The method is useful for promoting nerve regeneration and preventing neuronal degeneration in central/peripheral nervous system from injury/disease, where the injury is spinal cord injury, blunt trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or damages caused by surgery such as tumour excision. The disease is not an autoimmune disease or neoplasm. The disease results in a degenerative process occurring in either gray or white matter or both. The disease is diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea, amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and vitamin deficiency, intervertebral disc herniation, prion diseases such as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral neuropathies associated with various diseases, including but not limited to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary amyloidosis, obstructive lung diseases, acromegaly, malabsorption syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gammapathies, complications of various drugs (e.g., metronidazole) and toxins (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia

CC

```
CC
      telangectasia, Friedreich's ataxia, amyloid polyneuropathies,
 CC
     adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's
 CC
     disease, or lipoproteinemia. The present sequence represents the human
 CC
     neurotransmitter receptor protein Nogo-C, an example of NS-specific
CC
     antigen
XX
     Sequence 199 AA;
 SQ
   Query Match
                         71.3%; Score 503; DB 5; Length 199;
   Best Local Similarity
                         98.1%; Pred. No. 6.4e-51;
  Matches 102; Conservative
                                1; Mismatches
                                                 1; Indels
                                                               0; Gaps
                                                                           0;
Qу
           38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
              Db
            1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
           98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
QУ
              61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
Db
RESULT 11
AAY71559
     AAY71559 standard; protein; 199 AA.
XX
AC
     AAY71559;
XX
DT
     02-NOV-2000 (first entry)
XX
DΕ
     Rat Nogo C/Nogo A proteins derived fragment to construct mutant Nogo-C.
XX
KW
     Rat; neurite growth inhibitor; Nogo A; Nogo C; neural cell; myelin; CNS;
KW
     central nervous system; neoplastic disease; antiproliferative; glioma;
KW
     antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;
     degenerative nerve disease; Alzheimer's disease; Parkinson's disease;
KW
     hyperproliferative disorder; benign dysproliferative disorder; diagnosis;
KW
KW
     psoriasis; tissue hypertrophy; neuronal regeneration; treatment;
KW
     structural plasticity; screening; mutant; mutein.
XX
OS
     Rattus sp.
XX
FH
     Key
                    Location/Qualifiers
FT
    Region
                    1. .11
FT
                    /note= "Corresponds to residues 40-50 of rat Nogo C
FT
                    protein shown in AAY71312"
FT
    Region
                    12. .199
FT
                    /note= "Corresponds to residues 975-1162 of rat Nogo A
FT
                    protein shown in AAY71310"
XX
PN
    WO200031235-A2.
XX
PD
    02-JUN-2000.
XX
PF
    05-NOV-1999;
                   99WO-US026160.
XX
PR
    06-NOV-1998;
                   98US-0107446P.
XX
```

```
PA
      (SCHW/) SCHWAB M E.
 PΑ
      (CHEN/) CHEN M S.
 XX
 ΡI
      Schwab ME, Chen MS;
 XX
 DR
      WPI; 2000-400052/34.
 XX
     Nogo proteins and nucleic acids useful for treating neoplastic disorders
 PT
 PT
     of the central nervous system and inducing regeneration of neurons.
XX
PS
     Example; Page; 122pp; English.
XX
     The patent relates to neurite growth inhibitor Nogo which is free of all
CC
CC
     central nervous system (CNS) myelin material with which it is natively
CC
     associated. Nogo proteins and fragments displaying neurite growth
     inhibitory activity are used in the treatment of neoplastic disease of
CC
     the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma,
CC
CC
     ependyoma, pinealoma, haemangioblastoma, acoustic neuroma,
CC
     oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and
     degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases.
CC
     Therapeutics which promote Nogo activity can be used to treat or prevent
CC
CC
     hyperproliferative or benign dysproliferative disorders e.g. psoriasis
     and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be
CC
     used to inhibit production of Nogo protein to induce regeneration of
CC
     neurons or to promote structural plasticity of the CNS in disorders where
CC
     neurite growth, regeneration or maintenance are deficient or desired. The
CC
     animal models can be used in diagnostic and screening methods for
CC
     predisposition to disorders and to screen for or test molecules which can
CC
     treat or prevent disorders or diseases of the CNS. The present sequence
CC
     is derived by fusing two fragments from rat Nogo C and Nogo A proteins.
CC
CC
     The fragment is used in the construction of mutant Nogo-C which is
     composed of His-tag/T7-tag/Nogo-C N-terminus (11 aa) + Nogo-A sequence aa
CC
     975-1162. Nogo A deletion mutants were used for mapping the inhibitory
CC
     sites of Nogo protein. Major inhibitory region was identified in the Nogo
CC
     A sequence from amino acids 172-974, particularly amino acids 542-722. In
CC
     addition, N-terminal region 1-171 was found to be inhibitory to NIH 3T3
CC
     fibroblast spreading. Note: The present sequence is not given in the
CC
     specification but is derived from rat Nogo C sequence shown in AAY71312
CC
CC
     and Nogo A sequence shown in AAY71310. SEQ ID numbers 35-42 are referred
CC
     in claim 32 and SEQ ID NO: 29 in disclosure of the specification.
     However, the specification does not include sequences for these SEQ ID
CC
CC
     numbers
XX
SO
     Sequence 199 AA;
  Query Match
                         70.9%;
                                 Score 499.5; DB 3;
                                                     Length 199;
  Best Local Similarity
                         98.1%; Pred. No. 1.7e-50;
         103; Conservative
                                0; Mismatches
                                                 1;
                                                     Indels
                                                               1;
                                                                           1;
Qу
          38 MDGQKKHWKDK-VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTI 96
             1 MDGQKKHWKDKSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTI 60
Db
Qу
          97 SFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
          61 SFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 105
```

```
RESULT 12
 AAY71311
 ID
      AAY71311 standard; protein; 1178 AA.
 XX
 AC
      AAY71311;
 XX
 DT
      02-NOV-2000 (first entry)
 XX
 DΕ
      Human neurite growth inhibitor Nogo.
XX
 KW
      Human; neurite growth inhibitor; Nogo; neural cell; myelin; CNS;
      central nervous system; neoplastic disease; antiproliferative; glioma;
KW
KW
      antisense gene therapy; neuroblastoma; menagioma; retinoblastoma;
      degenerative nerve disease; Alzheimer's disease; Parkinson's disease;
KW
      hyperproliferative disorder; benign dysproliferative disorder; diagnosis;
KW
KW
      psoriasis; tissue hypertrophy; neuronal regeneration; treatment;
KW
      structural plasticity; screening.
XX
OS
     Homo sapiens.
XX
                      Location/Qualifiers
FH
FT
     Misc-difference 187
FT
                      /label= Unknown
FT
     Misc-difference 188
FT
                      /label= Unknown
     Misc-difference 189
FT
FT
                      /label= Unknown
FT
     Misc-difference 190
FT
                      /label= Unknown
FT
     Misc-difference 221
FT
                      /label= Unknown
FT
     Misc-difference 328
FT
                      /label= Unknown
FT
     Misc-difference 477
FT
                      /label= Unknown
FT
     Region
                      977. .1012
FT
                      /note= "Region specifically described in claim 16"
FT
     Region
FT
                      /note= "Region specifically described in claim 16"
FT
     Region
                      1079. .1114
FT
                      /note= "Region specifically described in claim 16"
XX
ΡN
     W0200031235-A2.
XX
PD
     02-JUN-2000.
XX
     05-NOV-1999;
PF
                    99WO-US026160.
XX
PR
     06-NOV-1998;
                    98US-0107446P.
XX
PA
     (SCHW/) SCHWAB M E.
PA
     (CHEN/) CHEN M S.
XX
PΙ
     Schwab ME,
                Chen MS;
XX
DR
     WPI; 2000-400052/34.
```

```
XX
      Nogo proteins and nucleic acids useful for treating neoplastic disorders
 PТ
      of the central nervous system and inducing regeneration of neurons.
 PT
 XX
 PS
      Claim 11; Fig 13; 122pp; English.
 XX
      The present sequence is a human Nogo protein which is a potent neural
 CC
      cell growth inhibitor and is free of all central nervous system (CNS)
 CC
      myelin material with which it is natively associated. The human Nogo
 CC
      sequence was derived by aligning human expressed sequence tags (ESTs)
 CC
      e.g. AA158636, AA333267, AA081783, AA167765, AA322918, AA092565, AA081525
 CC
      and AA081840 with the rat Nogo sequence. Nogo proteins and fragments
 CC
      displaying neurite growth inhibitory activity are used in the treatment
 CÇ
 CC
     of neoplastic disease of the CNS e.g. glioma, glioblastoma,
 CC
     medulloblastoma, craniopharyngioma, ependyoma, pinealoma,
 CC
     haemangioblastoma, acoustic neuroma, oligodendroglioma, menagioma,
     neuroblastoma or retinoblastoma and degenerative nerve diseases e.g.
CC
CC
     Alzheimer's and Parkinson's diseases. Therapeutics which promote Nogo
     activity can be used to treat or prevent hyperproliferative or benign
CC
     dysproliferative disorders e.g. psoriasis and tissue hypertrophy.
CC
CC
     Ribozymes or antisense Nogo nucleic acids can be used to inhibit
CC
     production of Nogo protein to induce regeneration of neurons or to
CC
     promote structural plasticity of the CNS in disorders where neurite
     growth, regeneration or maintenance are deficient or desired. The animal
CC
     models can be used in diagnostic and screening methods for predisposition
CC
     to disorders and to screen for or test molecules which can treat or
CC
     prevent disorders or diseases of the CNS. Note: SEQ ID numbers 35-42 are
CC
     referred in claim 32 and SEQ ID NO: 29 in disclosure of the
CC
     specification. However the specification does not include sequences for
CC
CC
     these SEQ ID numbers
XX
SQ
     Sequence 1178 AA;
  Query Match
                          63.5%; Score 448; DB 3; Length 1178;
  Best Local Similarity
                                Pred. No. 2.2e-43;
                         74.6%;
  Matches
           97; Conservative
                                3; Mismatches
                                                8; Indels
                                                               22;
                                                                    Gaps
                                                                            1:
           12 RENFAVYSVSVGMHNLLLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFL 71
Qу
                                               1 1111111111111111111111111
          976 RSPSAIFSADLG------KTSVVDLLYWRDIKKTGVVFGASLFL 1013
Db
           72 LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEEL 131
Qу
              1014 LLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEEL 1073
Db
Qy
          132 VQKYSNSALG 141
              11111111
Db
         1074 VQKYSNSALG 1083
RESULT 13
AAY71563
    AAY71563 standard; protein; 403 AA.
ID
XX
AC
    AAY71563;
XX
DT
    02-NOV-2000 (first entry)
```

XXDΕ Rat Nogo A protein fragment used in the construction of mutant EST. XX KW Rat; neurite growth inhibitor; Nogo A; neural cell; myelin; CNS; KW central nervous system; neoplastic disease; antiproliferative; glioma; antisense gene therapy; neuroblastoma; menagioma; retinoblastoma; KW KW degenerative nerve disease; Alzheimer's disease; Parkinson's disease; hyperproliferative disorder; benign dysproliferative disorder; diagnosis; KW psoriasis; tissue hypertrophy; neuronal regeneration; treatment; KW KW structural plasticity; screening; mutant; mutein. XX OS Rattus sp. XX PN WO200031235-A2. XX PD02-JUN-2000. XX PF05-NOV-1999; 99WO-US026160. XX 06-NOV-1998; PR 98US-0107446P. XX PA (SCHW/) SCHWAB M E. PA (CHEN/) CHEN M S. XX PΤ Schwab ME, Chen MS; XX DR

WPI; 2000-400052/34.

XX

РΤ PT

XX PS

XX

CC

CC CC

CC CC

CC CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC CC

CC

Nogo proteins and nucleic acids useful for treating neoplastic disorders of the central nervous system and inducing regeneration of neurons.

Example; Page; 122pp; English.

The patent relates to neurite growth inhibitor Nogo which is free of all central nervous system (CNS) myelin material with which it is natively associated. Nogo proteins and fragments displaying neurite growth inhibitory activity are used in the treatment of neoplastic disease of the CNS e.g. glioma, glioblastoma, medulloblastoma, craniopharyngioma, ependyoma, pinealoma, haemangioblastoma, acoustic neuroma, oligodendroglioma, menagioma, neuroblastoma or retinoblastoma and degenerative nerve diseases e.g. Alzheimer's and Parkinson's diseases. Therapeutics which promote Nogo activity can be used to treat or prevent hyperproliferative or benign dysproliferative disorders e.g. psoriasis and tissue hypertrophy. Ribozymes or antisense Nogo nucleic acids can be used to inhibit production of Nogo protein to induce regeneration of neurons or to promote structural plasticity of the CNS in disorders where neurite growth, regeneration or maintenance are deficient or desired. The animal models can be used in diagnostic and screening methods for predisposition to disorders and to screen for or test molecules which can treat or prevent disorders or diseases of the CNS. The present sequence is a fragment of rat Nogo A protein shown in AAY71310, which is used in the construction of mutant EST. The mutant is composed of His-tag/T7tag/Nogo-A sequence aa 760-1162. Nogo A deletion mutants were used for mapping the inhibitory sites of Nogo protein. Major inhibitory region was identified in the Nogo A sequence from amino acids 172-974, particularly amino acids 542-722. In addition, N-terminal region 1-171 was found to be inhibitory to NIH 3T3 fibroblast spreading. Note: The present sequence is

```
not given in the specification but is derived from rat Nogo A sequence
 CC
      shown in AAY71310. SEQ ID numbers 35-42 are referred in claim 32 and SEQ
 CC
      ID NO: 29 in disclosure of the specification. However, the specification
 CC
 CC
      does not include sequences for these SEQ ID numbers
 XX
      Sequence 403 AA;
 SQ
   Query Match
                           63.4%;
                                  Score 447; DB 3; Length 403;
   Best Local Similarity
                          96.9%;
                                   Pred. No. 6.9e-44;
   Matches
             93; Conservative
                                 0; Mismatches
                                                   3;
                                                                 0; Gaps
                                                       Indels
                                                                             0;
            46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
 Qу
                  214 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 273
 Db
 Qу
           106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
               111 1111111111111111111111111111111
 Db
           274 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 309
RESULT 14
AAY95012
     AAY95012 standard; protein; 893 AA.
ID
XX
AC
     AAY95012;
XX
DT
     19-JUN-2000 (first entry)
XX
DE
     Human secreted protein vb22_1, SEQ ID NO:64.
XX
     Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW
     blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW
KW
     infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW
     neurodegenerative disease; asthma; contraceptive.
XX
OS
     Homo sapiens.
XX
PN
     WO200011015-A1.
XX
PD
     02-MAR-2000.
XX
PF
     24-AUG-1999;
                    99WO-US019351.
XX
PR
     24-AUG-1998;
                    98US-0097638P.
     24-AUG-1998;
PR
                    98US-0097659P.
PR
     09-SEP-1998;
                    98US-0099618P.
PR
     28-SEP-1998;
                    98US-0102092P.
PR
     25-NOV-1998;
                    98US-0109978P.
PR
     23-DEC-1998;
                    98US-0113645P.
PR
     23-DEC-1998;
                    98US-0113646P.
PR
     23-AUG-1999;
                    99US-00379246.
XX
PA
     (ALPH-) ALPHAGENE INC.
XX
PI
    Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
DR
    WPI; 2000-224657/19.
```

```
XX
 PT
      New secreted or transmembrane proteins and polynucleotides encoding them,
      useful for treating neurodegenerative disorders, autoimmune diseases and
 PT
 PT
      cancer.
 XX
 PS
      Claim 73; Page 322-325; 357pp; English.
 XX
      The invention relates to 40 human secreted proteins (AAY94981-Y95020),
 CC
      and cDNA sequences encoding them (AAA23423-A23462). The secreted proteins
 CC
      of the invention include those that are thought to be only partially
 CC
      secreted, i.e., transmembrane proteins. The proteins of the invention may
 CC
      exhibit one or more activities selected from the following: cytokine
 CC
 CC
      activity; cell proliferation; differentiation; immune modulation;
 CC
     haematopoiesis regulation; tissue growth activity; activin/inhibin
 CC
      activity; chemotactic/chemokinetic activity; haemostatic and thrombolytic
     activity; anti-inflammatory activity; and tumour inhibition activity. The
 CC
     proteins may be administered to patients as vaccines, and the nucleotides
 CC
     may be used as part of a gene therapy regime. Diseases or conditions that
 CC
     may be treated using the proteins or nucleotides of the invention include
 CC
     autoimmune diseases; genetic disorders; haemophilia; cardiovascular
 CC
     diseases; cancer; bacterial, fungal and viral infections, especially HIV;
 CC
 CC
     multiple sclerosis; rheumatoid arthritis; pulmonary inflammation;
     Guillain-Barre syndrome; insulin dependent diabetes mellitus; and
 CC
     allergic reactions such as asthma and anaemia. They may also be used for
 CC
     treating wounds, burns, ulcers, osteoporosis, osteoarthritis, periodontal
 CC
     diseases, Alzheimer's disease, Parkinson's disease, Huntington's disease
CC
     and amyotrophic lateral sclerosis (ALS). Proteins with activin/inhibin
CC
     activity may additionally be useful as contraceptives. Nucleic acid
CC
     sequences of the invention may be used in chromosome mapping, and as a
CC
     source of diagnostic primers and probes. The present sequence represents
CC
CC
     one of the 40 proteins of the invention
XX
SQ
     Sequence 893 AA;
  Query Match
                                 Score 447; DB 3; Length 893;
                          63.4%;
  Best Local Similarity
                          96.9%;
                                 Pred. No. 2e-43;
           93; Conservative
                                0; Mismatches
                                                  3; Indels
                                                               0;
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
                703 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 762
Db
          106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
              Db
          763 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 798
RESULT 15
     ABU11573 standard; protein; 983 AA.
XX
     ABU11573;
AC
XX
DT
     12-FEB-2003
                 (first entry)
XX
DΕ
     Human MDDT polypeptide SEQ ID 520.
XX
```

MDDT; human; disease detection and treatment molecule polypeptide; KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV; KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic; KW gene therapy; protein replacement therapy; cell proliferative disorder; KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma; KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS; KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia; KW KW psoriasis; hepatitis. XX OS Homo sapiens. XX PN WO200279449-A2. XX PD10-OCT-2002. XX PF27-MAR-2002; 2002WO-US009944. XX PR 28-MAR-2001; 2001US-0279619P. PR 29-MAR-2001; 2001US-0280067P. PR29-MAR-2001; 2001US-0280068P. 16-MAY-2001; 2001US-0291280P. PR PR 17-MAY-2001; 2001US-0291829P. PR 17-MAY-2001; 2001US-0291849P. PR 19-JUN-2001; 2001US-0299428P. PR 20-JUN-2001; 2001US-0299776P. PR 20-JUN-2001; 2001US-0300001P. XX PΑ (INCY-) INCYTE GENOMICS INC. XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J; PΙ Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR; PΙ Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH; ΡI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B; PΙ PΙ Flores V, Marwaha R, Lo A, Lan RY, Urashka ME; XX DR WPI; 2003-058431/05. DR N-PSDB; ABX34563. XX New purified disease detection and treatment molecule proteins and PTpolynucleotides, useful for diagnosing, treating or preventing cancers PT(e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis PTPTor hepatitis. XX Claim 27; SEQ ID NO 520; 339pp + Sequence Listing; English. PS XX This invention describes a novel disease detection and treatment molecule CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive, CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic, CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides CC and the polypeptides of the invention can be used for gene therapy, CC protein replacement therapy and are useful for treating a variety of CC diseases or conditions. These polypeptides or polynucleotides are CC particularly useful for diagnosing, treating or preventing cell CCproliferative disorders (e.g. cancers including adenocarcinoma, CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's CCsyndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or CC

```
hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded
    by ABU11450-ABU11845, described in the disclosure of the invention. NOTE:
CC
    The sequence data for this patent did not form part of the printed
CC
CC
    specification, but was obtained in electronic format from WIPO at
CC
    ftp.wipo.int/pub/published_pct_sequences
XX
    Sequence 983 AA;
SQ
 Query Match
                       63.4%; Score 447; DB 6; Length 983;
 Best Local Similarity
                      96.9%; Pred. No. 2.3e-43;
         93; Conservative 0; Mismatches
                                           3; Indels
                                                         0; Gaps
                                                                   0;
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qy
            793 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 852
Db
        106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
            Db
        853 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 888
```

Search completed: September 3, 2004, 16:05:33 Job time : 12.4871 secs

CC

## GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2004, 16:02:30; Search time 2.78364 Seconds

(without alignments)

2615.013 Million cell updates/sec

Title:

US-09-830-972-32

Perfect score:

e: 705

Sequence:

1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVQKYSNSALG 141

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched:

389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:

389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgn2 6/ptodata/2/iaa/5A COMB.pep:\*
- 2: /cgn2 6/ptodata/2/iaa/5B COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2 6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		8				
Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	503	71.3	199	2	US-08-700-607-1	Sequence 1, Appli
2	349	49.5	208	2	US-08-700-607-7	Sequence 7, Appli
3	348	49.4	267	2	US-08-700-607-8	Sequence 8, Appli
4	337	47.8	356	2	US-08-700-607-6	Sequence 6, Appli
5	337	47.8	776	2	US-08-700-607-5	Sequence 5, Appli
6	305	43.3	241	2	US-08-700-607-3	Sequence 3, Appli
7	227	32.2	168	4	US-09-149-476-563	Sequence 563, App
8	99	14.0	80	3	US-08-905-223-411	Sequence 411, App
9	75	10.6	593	4	US-09-328-352-4866	Sequence 4866, Ap
10	72.5	10.3	598	2	US-08-853-659A-53	Sequence 53, Appl
11	71.5	10.1	154	1	US-08-366-783-5	Sequence 5, Appli
						1

```
12
        70
              9.9
                     518 4 US-09-134-001C-4744
                                                       Sequence 4744, Ap
13
        70
              9.9
                     563 4 US-09-422-936-79
                                                       Sequence 79, Appl
14
        70
              9.9
                     619 3 US-08-262-220-6
                                                       Sequence 6, Appli
15
        70
              9.9
                     619 3 US-08-471-733-6
                                                       Sequence 6, Appli
16
        70
              9.9
                     619 3 US-08-468-878-6
                                                       Sequence 6, Appli
17
        70
              9.9
                     619 3 US-08-750-494-6
                                                       Sequence 6, Appli
18
        70
              9.9
                     619. 4 US-08-470-638-6
                                                       Sequence 6, Appli
19
        70
              9.9
                     844 4 US-09-422-936-47
                                                       Sequence 47, Appl
20
        70
              9.9
                    844 4 US-09-422-936-51
                                                       Sequence 51, Appl
21
        70
              9.9
                    886 4 US-09-422-936-77
                                                       Sequence 77, Appl
22
        70
             9.9
                    892 4 US-09-422-936-75
                                                       Sequence 75, Appl
23
        70
             9.9
                    899 4 US-09-422-936-71
                                                       Sequence 71, Appl
24
        70
             9.9
                    960 4 US-09-422-936-45
                                                       Sequence 45, Appl
25
        70
             9.9
                    961
                         4 US-09-422-936-49
                                                       Sequence 49, Appl
26
        70
             9.9
                    961 4 US-09-914-259-14
                                                       Sequence 14, Appl
27
        69
             9.8
                    621 3 US-08-262-220-8
                                                       Sequence 8, Appli
28
        69
             9.8
                    621 3 US-08-471-733-8
                                                       Sequence 8, Appli
29
        69
                    621 3 US-08-468-878-8
             9.8
                                                       Sequence 8, Appli
30
        69
                    621 3 US-08-750-494-8
             9.8
                                                      Sequence 8, Appli
31
        69
                    621 4 US-08-470-638-8
             9.8
                                                      Sequence 8, Appli
32
        68
             9.6
                    344 4 US-09-107-532A-6886
                                                      Sequence 6886, Ap
33
       68
             9.6
                   1447 3 US-09-041-886-25
                                                      Sequence 25, Appl
34
       68
                   1447 5 PCT-US94-05277-2
             9.6
                                                      Sequence 2, Appli
35
     67.5
             9.6
                   659 4 US-09-328-352-6021
                                                      Sequence 6021, Ap
36
       67
             9.5
                    231 4
                           US-09-198-452A-419
                                                      Sequence 419, App
37
       67
             9.5
                    493 4
                           US-09-540-236-2120
                                                      Sequence 2120, Ap
38
       67
             9.5
                    507 4
                            US-09-328-352-7742
                                                      Sequence 7742, Ap
39
      66.5
             9.4
                            US-09-107-532A-6706
                    249
                        4
                                                      Sequence 6706, Ap
40
       66
             9.4
                    445
                        4
                            US-09-328-352-4714
                                                      Sequence 4714, Ap
41
      65.5
             9.3
                    187
                         2
                            US-08-846-021A-5
                                                      Sequence 5, Appli
42
     65.5
             9.3
                    234 1
                            US-08-366-783-4
                                                      Sequence 4, Appli
43
     65.5
             9.3
                    254 2
                            US-08-767-026-7
                                                      Sequence 7, Appli
44
     65.5
             9.3
                    254 4
                           US-09-319-275A-7
                                                      Sequence 7, Appli
45
       65
             9.2
                    424 4 US-09-543-681A-7510
                                                      Sequence 7510, Ap
```

## ALIGNMENTS

```
RESULT 1
US-08-700-607-1
; Sequence 1, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
     APPLICANT: Bandman, Olga
     APPLICANT: Au-Young, Janice
;
     APPLICANT: Goli, Surya K.
     APPLICANT: Hillman, Jennifer L.
     TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
;
       ADDRESSEE: Incyte Pharmaceuticals, Inc.
;
      STREET: 3174 Porter Drive
      CITY: Palo Alto
;
      STATE: CA
      COUNTRY: U.S.
;
      ZIP: 94304
```

```
COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
       NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
       REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: 415-855-0555
       TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO: 1:
     SEQUENCE CHARACTERISTICS:
      LENGTH: 199 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
    IMMEDIATE SOURCE:
      LIBRARY:
      CLONE: Consensus
US-08-700-607-1
  Query Match
                        71.3%; Score 503; DB 2; Length 199;
  Best Local Similarity 98.1%; Pred. No. 2.9e-52;
 Matches 102; Conservative 1; Mismatches 1; Indels
                                                            0; Gaps
                                                                       0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
             1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Db
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 2
US-08-700-607-7
; Sequence 7, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Au-Young, Janice
    APPLICANT: Goli, Surya K.
    APPLICANT: Hillman, Jennifer L.
    TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Incyte Pharmaceuticals, Inc.
      STREET: 3174 Porter Drive
      CITY: Palo Alto
      STATE: CA
      COUNTRY: U.S.
```

```
ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
      NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
       REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: 415-855-0555
       TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO: 7:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 208 amino acids
       TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
    IMMEDIATE SOURCE:
      LIBRARY: GenBank
      CLONE: 307311
US-08-700-607-7
  Query Match
                        49.5%; Score 349; DB 2; Length 208;
  Best Local Similarity 63.4%; Pred. No. 8.6e-34;
 Matches
           64; Conservative 19; Mismatches 18; Indels
                                                              0; Gaps
                                                                         0;
          37 EMDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTI 96
Qу
                    Db
           9 KMDCVWSNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATI 68
          97 SFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
Qу
             111111 1:11: 1:111111:1111 1: :1:1 :111::
Db
          69 SFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 109
RESULT 3
US-08-700-607-8
; Sequence 8, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Au-Young, Janice
    APPLICANT: Goli, Surya K.
    APPLICANT: Hillman, Jennifer L.
    TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
     ADDRESSEE: Incyte Pharmaceuticals, Inc.
     STREET: 3174 Porter Drive
     CITY: Palo Alto
     STATE: CA
```

```
COUNTRY: U.S.
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
       NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
       REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: 415-855-0555
       TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO:
     SEQUENCE CHARACTERISTICS:
      LENGTH: 267 amino acids
;
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
    IMMEDIATE SOURCE:
      LIBRARY: GenBank
      CLONE: 281046
US-08-700-607-8
  Query Match
                        49.4%; Score 348; DB 2; Length 267;
  Best Local Similarity 64.0%; Pred. No. 1.6e-33;
 Matches
         64; Conservative 18; Mismatches 18; Indels
                                                           0; Gaps
                                                                       0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
                  Db
           1 MDCVWSNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATIS 60
Qу
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
             Db
          61 FRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 100
RESULT 4
US-08-700-607-6
; Sequence 6, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Au-Young, Janice
    APPLICANT: Goli, Surya K.
    APPLICANT: Hillman, Jennifer L.
    TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Incyte Pharmaceuticals, Inc.
      STREET: 3174 Porter Drive
      CITY: Palo Alto
```

```
STATE: CA
       COUNTRY: U.S.
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
       NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
       REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: 415-855-0555
       TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 356 amino acids
       TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
     IMMEDIATE SOURCE:
      LIBRARY: GenBank
      CLONE: 307309
US-08-700-607-6
  Query Match
                        47.8%; Score 337; DB 2; Length 356;
  Best Local Similarity 67.4%; Pred. No. 5e-32;
           62; Conservative 16; Mismatches 14; Indels
 Matches
                                                             0; Gaps
                                                                         0;
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qy
             166 KQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 225
Db
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
             11: 1:11111:111 1: :1:1 :111::
Db
         226 QAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 257
RESULT 5
US-08-700-607-5
; Sequence 5, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Au-Young, Janice
    APPLICANT: Goli, Surya K.
    APPLICANT: Hillman, Jennifer L.
    TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Incyte Pharmaceuticals, Inc.
      STREET: 3174 Porter Drive
```

```
CITY: Palo Alto
       STATE: CA
       COUNTRY: U.S.
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
       NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
       REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-855-0555
      TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 776 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
    IMMEDIATE SOURCE:
      LIBRARY: GenBank
      CLONE: 307307
US-08-700-607-5
 Query Match
                        47.8%; Score 337; DB 2; Length 776;
 Best Local Similarity
                       67.4%; Pred. No. 1.5e-31;
          62; Conservative 16; Mismatches 14; Indels
 Matches
                                                             0; Gaps
                                                                         0;
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
             586 KQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 645
Db
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
Qу
             11: 1:111111:111 1: :1:1 :111::
Db
         646 QAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 677
RESULT 6
US-08-700-607-3
; Sequence 3, Application US/08700607
; Patent No. 5858708
  GENERAL INFORMATION:
    APPLICANT: Bandman, Olga
    APPLICANT: Au-Young, Janice
    APPLICANT: Goli, Surya K.
    APPLICANT: Hillman, Jennifer L.
    TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
    NUMBER OF SEQUENCES: 9
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```
STREET: 3174 Porter Drive
       CITY: Palo Alto
       STATE: CA
       COUNTRY: U.S.
       ZIP: 94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE: Diskette
       COMPUTER: IBM Compatible
       OPERATING SYSTEM: DOS
       SOFTWARE: FastSEQ Version 1.5
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/700,607
       FILING DATE: Filed Herewith
     ATTORNEY/AGENT INFORMATION:
      NAME: Billings, Lucy J.
       REGISTRATION NUMBER: 36,749
      REFERENCE/DOCKET NUMBER: PF-0114 US
     TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-855-0555
      TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 241 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
     IMMEDIATE SOURCE:
      LIBRARY: THP1NOB01
      CLONE: 31870
US-08-700-607-3
  Query Match
                        43.3%; Score 305; DB 2; Length 241;
  Best Local Similarity 60.9%; Pred. No. 2e-28;
  Matches 56; Conservative 17; Mismatches 19; Indels
                                                           0; Gaps
          49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
Qу
             48 VHDLIXWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAV 107
Db
         109 AKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
Qу
              Db
         108 QKSEEGHPFKAYLDVDITLSSEAFHNYMNAAM 139
RESULT 7
US-09-149-476-563
; Sequence 563, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
  APPLICANT: Rosen et al.
  TITLE OF INVENTION: 186 Human Secreted proteins
  FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
```

```
; EARLIER APPLICATION NUMBER: 60/040,162
```

- EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/040,333
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/038,621
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/040,626
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/040,334
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/040,336
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/040,163
- ; EARLIER FILING DATE: 1997-03-07
- ; EARLIER APPLICATION NUMBER: 60/047,600
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,615
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,597
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,502
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,633
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,583
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,617
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,618
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,503
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,592
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,581
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,584
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,500
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,587
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,492
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,598
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,613
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,582
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,596
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,612
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,632
- ; EARLIER FILING DATE: 1997-05-23
- EARLIER APPLICATION NUMBER: 60/047,601

```
; EARLIER FILING DATE: 1997-05-23
```

- EARLIER APPLICATION NUMBER: 60/043,580
- EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,568
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,314
- ; EARLIER FILING DATE: 1997-04-11
- EARLIER APPLICATION NUMBER: 60/043,569
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,311
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,671
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,674
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,669
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,312
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,313
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,672
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,315
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/048,974
- ; EARLIER FILING DATE: 1997-06-06
- ; EARLIER APPLICATION NUMBER: 60/056,886
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,877
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,889
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,893
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,630
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,878
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,662
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,872
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,882
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,637
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,903
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,888
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,879
- EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,880
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,894
- ; EARLIER FILING DATE: 1997-08-22

```
; EARLIER APPLICATION NUMBER: 60/056,911
```

- ; EARLIER FILING DATE: 1997-08-22
- EARLIER APPLICATION NUMBER: 60/056,636
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,874
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,910
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,864
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,631
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,845
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,892
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/057,761
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/047,595
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,599
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,588
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,585
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,586
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,590
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,594
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,589
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,593
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/047,614
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/043,578
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/043,576
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/047,501
- ; EARLIER FILING DATE: 1997-05-23
- ; EARLIER APPLICATION NUMBER: 60/043,670
- ; EARLIER FILING DATE: 1997-04-11
- ; EARLIER APPLICATION NUMBER: 60/056,632
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,664
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,876
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,881
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,909
- ; EARLIER FILING DATE: 1997-08-22
- ; EARLIER APPLICATION NUMBER: 60/056,875

```
EARLIER FILING DATE: 1997-08-22
   EARLIER APPLICATION NUMBER: 60/056,862
   EARLIER FILING DATE: 1997-08-22
   EARLIER APPLICATION NUMBER: 60/056,887
   EARLIER FILING DATE: 1997-08-22
   EARLIER APPLICATION NUMBER: 60/056,908
   EARLIER FILING DATE: 1997-08-22
   EARLIER APPLICATION NUMBER: 60/048,964
   EARLIER FILING DATE: 1997-06-06
   EARLIER APPLICATION NUMBER: 60/057,650
   EARLIER FILING DATE: 1997-09-05
   EARLIER APPLICATION NUMBER: 60/056,884
   EARLIER FILING DATE: 1997-08-22
   EARLIER APPLICATION NUMBER: 60/057,669
   EARLIER FILING DATE: 1997-09-05
   EARLIER APPLICATION NUMBER: 60/049,610
  EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
  Query Match
                         32.2%; Score 227; DB 4; Length 168;
  Best Local Similarity
                         60.0%; Pred. No. 2.7e-19;
  Matches
           42; Conservative 14; Mismatches
                                              14; Indels
                                                               0;
                                                                  Gaps
          71 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEE 130
Qy
             Db
           1 MLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSE 60
         131 LVQKYSNSAL 140
Qy
                 1 |:|:
Db
          61 AFHNYMNAAM 70
RESULT 8
US-08-905-223-411
; Sequence 411, Application US/08905223
; Patent No. 6222029
  GENERAL INFORMATION:
    APPLICANT: Edwards, Jean-Baptiste D.
    APPLICANT: Duelert, Aymeric
    APPLICANT: Lacroix, Bruno
    TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
    NUMBER OF SEQUENCES: 503
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: Knobbe, Martens, Olson & Bear
      STREET: 501 West Broadway
      CITY: San Diego
      STATE: California
      COUNTRY: USA
      ZIP: 92101-3505
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy Disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: Win95
      SOFTWARE: Word
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/905,223
```

```
FILING DATE:
        CLASSIFICATION: 536
     ATTORNEY/AGENT INFORMATION:
      NAME: Israelsen, Ned A.
       REGISTRATION NUMBER: 29,655
       REFERENCE/DOCKET NUMBER:
     TELECOMMUNICATION INFORMATION:
       TELEPHONE: (619) 235-8550
       TELEFAX: (619) 235-0176
   INFORMATION FOR SEQ ID NO: 411:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 80 amino acids
       TYPE: AMINO ACID
       TOPOLOGY: LINEAR
     MOLECULE TYPE: PROTEIN
     ORIGINAL SOURCE:
       ORGANISM: Homo Sapiens
       TISSUE TYPE: Brain
    FEATURE:
      NAME/KEY: sig_peptide
       LOCATION: -78..-1
       IDENTIFICATION METHOD: Von Heijne matrix
       OTHER INFORMATION: score 6.3
       OTHER INFORMATION: seq TLIMLLSWQLSVS/SV
US-08-905-223-411
  Query Match
                          14.0%; Score 99; DB 3; Length 80;
  Best Local Similarity 58.8%; Pred. No. 0.0002;
  Matches 20; Conservative 7; Mismatches 5; Indels 2; Gaps
                                                                             1;
Qy
           49 VVDLLYWRDIKKTGVVFGASLFLLLS--LTVFSI 80
              1 11::111:111 111 :1 :111 1:1 1:
Db
           47 VHDLIFWRDVKKTGFVFGTTLIMLLSWQLSVSSV 80
RESULT 9
US-09-328-352-4866
; Sequence 4866, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
  FILE REFERENCE: GTC99-03PA
  CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4866
   LENGTH: 593
   TYPE: PRT
   ORGANISM: Acinetobacter baumannii
US-09-328-352-4866
 Query Match 10.6%; Score 75; DB 4; Length 593; Best Local Similarity 27.5%; Pred. No. 2.5;
 Matches 25; Conservative 19; Mismatches 27; Indels 20; Gaps
```

```
Qу
           55 WRDIKKT--GVVFGASLFLLLSLTVFSIV----SVTAYIALALLSVTISFRIYKGVIQA 107
             258 WRGVKETWPAVLVGGGAFAIAQYLTSNFIGPELPDITAAIA-SLVSLTLLFRVWK---- 311
 Db
         108 IAKSDEGHPFRAYLESEVAISEE--LVQKYS 136
 Qу
                 Db
         312 ----PKHIFRFEPEAGQTLAQQPTTVQRYS 337
RESULT 10
US-08-853-659A-53
 ; Sequence 53, Application US/08853659A
 ; Patent No. 5925522
   GENERAL INFORMATION:
     APPLICANT: Wong, K.K.; Saffer, J.D.
     TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
     TITLE OF INVENTION: Of A
     TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of
Salmonella
     NUMBER OF SEQUENCES: 67
     CORRESPONDENCE ADDRESS:
      ADDRESSEE: Paul W. Zimmerman
;
      ADDRESSEE: Intellectual Property Services
;
      ADDRESSEE: Battelle Memorial Institute
     ADDRESSEE: PNNL P.O. Box 999
     STREET: Washington Way
      CITY: Richland
     STATE: Washington
      COUNTRY: U.S.A.
      ZIP: 99352
    COMPUTER READABLE FORM:
;
     MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
      COMPUTER: IBM PC/XT/AT
     OPERATING SYSTEM: MS-DOS
      SOFTWARE: Word Processor (WordPerfect 5.1)
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/853,659A
      FILING DATE: Unknown
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: none
      FILING DATE: n/a
  INFORMATION FOR SEQ ID NO: 53:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 598 amino acids
;
      TYPE: amino acid
      TOPOLOGY: linear
US-08-853-659A-53
 Query Match
                       10.3%; Score 72.5; DB 2; Length 598;
 Best Local Similarity 22.5%; Pred. No. 5;
 Matches 25; Conservative 22; Mismatches 53; Indels 11; Gaps 2;
QУ
         41 QKKHWKDKVVDLLYWRDIKKTGVVFGASLF-LLLSLTVFSIVSVTAYIALAL----- 91
            149 KKKNGRSMSSAFVLWNEFQKIKPVLLNSIFQRIADIPIFIIFLIVIYVNLGLVVIVPITM 208
```

```
Qу
           92 --LSVTISFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
               209 FIVSIIISLVNHHYTNELMNKQKEGQKNRNIFISEVFLSIKMIHTLNNQGL 259
 Db
 RESULT 11
 US-08-366-783-5
 ; Sequence 5, Application US/08366783
 ; Patent No. 5650554
   GENERAL INFORMATION:
     APPLICANT: Moloney, Maurice M
     TITLE OF INVENTION: Oil-Body Proteins As Carriers Of
     TITLE OF INVENTION: High-Value Peptides In Plants
     NUMBER OF SEQUENCES: 22
     CORRESPONDENCE ADDRESS:
       ADDRESSEE: DEHLINGER & ASSOCIATES
      STREET: 350 CAMBRIDGE AVENUE, SUITE 250
     CITY: PALO ALTO
      STATE: California
     COUNTRY: United States
     ZIP: 94025-1536
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/366,783
     FILING DATE:
      CLASSIFICATION: 800
    ATTORNEY/AGENT INFORMATION:
     NAME: FABIAN, GARY
      REGISTRATION NUMBER: 33,875
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-324-0880
      TELEFAX: 415-324-0960
   INFORMATION FOR SEQ ID NO: 5:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 154 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: peptide
US-08-366-783-5
  Query Match
                       10.1%; Score 71.5; DB 1; Length 154;
  Best Local Similarity 26.2%; Pred. No. 1;
         37; Conservative 16; Mismatches 49; Indels 39; Gaps
          33 RSWQEMDGQKKHWKDKVVDLLYWRDIKK--TGVVFGASLFLLLSLT------ 76
Qу
            6 RDQYQMSGRGSDYSKS-----RQIAKAATAVTAGGSLLVLLSLTLVGTVIALTVATPL 58
Db
QУ
         77 --VFSIVSVTAYIALALL------SVTISFRIYKG-VIQAIAKSDEGHPFRAY 120
              : | | : | | : | |
                                        ::|: ||| :|:
         59 LVIFSPILVPALITVALLITGFLSSGGFGIAAITVFSWIYKYLLIEHPQGSDKLDSARMK 118
Db
```

```
Qу
           121 LESEVAISEELVQKYSNSALG 141
               1 1: :: 1 1
                                1
 Db
           119 LGSKAQDLKDRAQYYGQQHTG 139
 RESULT 12
 US-09-134-001C-4744
 ; Sequence 4744, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
   APPLICANT: Lynn Doucette-Stamm et al
   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 STAPHYLOCOCCUS
   TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
   FILE REFERENCE: GTC-007
   CURRENT APPLICATION NUMBER: US/09/134,001C
   CURRENT FILING DATE: 1998-08-13
   PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4744
    LENGTH: 518
    TYPE: PRT
    ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4744
  Query Match
                          9.9%; Score 70; DB 4; Length 518;
  Best Local Similarity 21.0%; Pred. No. 8.1;
          21; Conservative 28; Mismatches 35; Indels
  Matches
                                                            16; Gaps
                                                                          4;
Qy
           16 AVYSVSVGMH---NLLLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKK----TGVVFGAS 68
              1: :::: |
                         400 AIVAITIAWHPNDTILNLVGNAWA---GFGAAFSPLVLYSLYWKDLTRAGAISGMVAGAV 456
Db
Qу
           69 LFLLLSLTVFSIVSVTAYIAL-----ALLSVTISFRIYK 102
                    : ::: |: :
                                        :: | | | :: : |
         457 VVIVWISWIKPLATINAFFGMYEIIPGFIVSVLITYIVSK 496
Db
RESULT 13
US-09-422-936-79
; Sequence 79, Application US/09422936
; Patent No. 6465213
; GENERAL INFORMATION:
  APPLICANT: Ekstrand, Jonas
  TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES
  FILE REFERENCE: 06275-165002
  CURRENT APPLICATION NUMBER: US/09/422,936
  CURRENT FILING DATE: 1999-10-22
  PRIOR APPLICATION NUMBER: US 09/242,608
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: PCT/SE98/01947
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9703914-2
```

```
PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9800864-2
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: SWEDEN 9802575-2
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 85
 SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 79
   LENGTH: 563
   TYPE: PRT
   ORGANISM: Homo sapiens
US-09-422-936-79
  Query Match
                         9.9%; Score 70; DB 4; Length 563;
  Best Local Similarity 28.3%; Pred. No. 9.1;
  Matches 26; Conservative 13; Mismatches
                                            35; Indels 18; Gaps
                                                                       3;
          30 LEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKTGVVFGAS-----LFLLLSLTVFS 79
Qу
             144 LQGGSYKKIG----YYDSTKDDLSW---SKTDKWIGGSPPADQTLVIKTFRFLSQKLFI 195
Db
Qу
         80 IVSVTAYIALALLSVTISFRIYKGVIQAIAKS 111
             Db
         196 SVSVLSSLGIVLAVVCLSFNIYNSHVRYIQNS 227
RESULT 14
US-08-262-220-6
; Sequence 6, Application US/08262220
; Patent No. 6054296
 GENERAL INFORMATION:
    APPLICANT: BERGSTROM SVEN
    APPLICANT: BARBOUR ALAN G.
    TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
    NUMBER OF SEQUENCES: 14
    CORRESPONDENCE ADDRESS:
     ADDRESSEE: BROWDY AND NEIMARK
      STREET: 419 SEVENTH STREET, N.W.
     CITY: WASHINGTON
     COUNTRY: USA
     ZIP: 20004
    COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/262,220
     FILING DATE: 20-JUN-1994
     CLASSIFICATION: 514
   ATTORNEY/AGENT INFORMATION:
     NAME: COOPER, IVER P.
     REGISTRATION NUMBER: 28,005
     REFERENCE/DOCKET NUMBER: BERGSTROM=3
   TELECOMMUNICATION INFORMATION:
     TELEPHONE: 202-628-5197
     TELEFAX: 202-737-3528
```

```
TELEX: 248633
   INFORMATION FOR SEQ ID NO: 6:
     SEQUENCE CHARACTERISTICS:
       LENGTH: 619 amino acids
       TYPE: amino acid
       TOPOLOGY: linear
     MOLECULE TYPE: protein
US-08-262-220-6
  Query Match
                          9.9%; Score 70; DB 3; Length 619;
  Best Local Similarity 39.0%; Pred. No. 10;
          16; Conservative 6; Mismatches 15; Indels 4; Gaps
  Matches
                                                                          1;
          32 GRSWQEMDGQKKHWKDKVVD----LLYWRDIKKTGVVFGAS 68
Qу
             1:1:1:1:1:1
                                    |:
                                           Db
         384 GLAWNKDDGEKESWKVKGSDSYSTRLFGEQDKKSGVALGIS 424
RESULT 15
US-08-471-733-6
; Sequence 6, Application US/08471733
; Patent No. 6068842
; GENERAL INFORMATION:
    APPLICANT: BERGSTROM SVEN
    APPLICANT: BARBOUR ALAN G.
    TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
    NUMBER OF SEQUENCES: 14
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: BROWDY AND NEIMARK
      STREET: 419 SEVENTH STREET, N.W.
     CITY: WASHINGTON
     COUNTRY: USA
      ZIP: 20004
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/471,733
      FILING DATE: 06-JUN-1995
      CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
     APPLICATION NUMBER: 08/262,220
      FILING DATE: 20-JUN-1994
   ATTORNEY/AGENT INFORMATION:
     NAME: COOPER, IVER P.
     REGISTRATION NUMBER: 28,005
     REFERENCE/DOCKET NUMBER: BERGSTROM=3
   TELECOMMUNICATION INFORMATION:
     TELEPHONE: 202-628-5197
     TELEFAX: 202-737-3528
     TELEX: 248633
INFORMATION FOR SEQ ID NO: 6:
   SEQUENCE CHARACTERISTICS:
     LENGTH: 619 amino acids
     TYPE: amino acid
```

Search completed: September 3, 2004, 16:10:40 Job time: 4.78364 secs

# GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2004, 16:00:14; Search time 3.01088 Seconds

(without alignments)

4504.667 Million cell updates/sec

Title: US-

US-09-830-972-32

Perfect score: 705

Sequence: 1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVQKYSNSALG 141

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1 2 3 4 5 6 7 8 9 10 11 12 13	349 348 337 162 159 158 83.5 81.5 79 78.5 77.5 77.5	49.5 49.4 47.8 23.0 22.6 22.4 11.8 11.6 11.2 11.1 11.0	208 267 776 2484 2607 222 295 464 618 458 261 583 481	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	I60904 A60021 A46583 T26216 T26215 T26213 S59439 C88188 T24228 A72258 F64924 T49359 C95920	neuroendocrine-spe tropomyosin-relate neuroendocrine-spe hypothetical prote hypothetical prote hypothetical prote probable membrane protein C18H9.5 [i hypothetical prote hypothetical prote probable thiosulfa hypothetical prote hypothetical membr

14	75.5	10.7	545	2	F64665
15	74.5	10.6	545	2	E71851
16	74	10.5	268	2	F64024
17	73	10.4	393	2	S67763
18	73	10.4	888	2	T01081
19	72.5	10.3	278	2	AD0147
20	72.5	10.3	302	2	AE2863
21	72.5	10.3	302	2	D97640
22	72.5	10.3	417	2	B96977
23	72.5	10.3	598	2	T14886
24	72	10.2	271	2	T13013
25	72	10.2	299	2	B69155
26	72	10.2	1783	2	T42386
27	72	10.2	3511	2	A59295
28	71.5	10.1	255	2	E84899
29	71.5	10.1	537	2	G82873
30	71	10.1	346	1	WMVZ1W
31	71	10.1	346	1	WMVZ2W
32	71	10.1	346	2	T37430
33	71	10.1	476	2	B97096
34	71	10.1	1065	2	T25068
35	70.5	10.0	153	2	AB3226
36	70.5	10.0	261	2	A90926
37	70.5	10.0	261	2	E85774
38	70.5	10.0	291	2	B69098
39	70	9.9	271	2	AC1320
40	70	9.9	619	2	S55502
41	70	9.9	737	2	AE1678
42	69.5	9.9	188	2	T04714
43	69.5	9.9	403	2	T04821
44	69.5	9.9	415	2	D95248
45	69.5	9.9	504	2	E83898

glucose-6-phosphat glucose-6-phosphat hypothetical prote probable membrane hypothetical prote probable ABC trans conserved hypothet hypothetical prote probable Mn transp leukotoxin express hypothetical prote hypothetical prote unconventional myo unconventional myo hypothetical prote conserved hypothet 3beta-hydroxy-Delt 3beta-hydroxy-Delt hydroyxsteroid deh 2-oxoglutarate/mal hypothetical prote conserved hypothet hypothetical prote hypothetical prote phosphate transpor hypothetical prote membrane-associate heavy metal-transp hypothetical prote hypothetical prote conserved hypothet spore germination

#### ALIGNMENTS

### RESULT 1 160904

neuroendocrine-specific protein C - human

C; Species: Homo sapiens (man)

C; Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 05-Nov-1999

C; Accession: I60904

R; Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.

J. Biol. Chem. 268, 13439-13447, 1993

A; Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.

A; Reference number: A46583; MUID: 93293865; PMID: 7685762

A; Accession: 160904

A; Status: preliminary; translated from GB/EMBL/DDBJ

A; Molecule type: mRNA A; Residues: 1-208 < RES>

A;Cross-references: GB:L10335; NID:g307310; PIDN:AAA59952.1; PID:g307311

C; Genetics:

A; Gene: GDB: RTN1; NSP

```
Query Match
                          49.5%; Score 349; DB 2; Length 208;
   Best Local Similarity
                         63.4%; Pred. No. 5.3e-28;
           64; Conservative 19; Mismatches
                                               18;
                                                    Indels
                                                              0; Gaps
                                                                         0:
           37 EMDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTI 96
 Qv
                    9 KMDCVWSNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATI 68
 Db
 Qу
           97 SFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
              Db
           69 SFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 109
 RESULT 2
 A60021
tropomyosin-related protein, neuronal - rat
 C; Species: Rattus norvegicus (Norway rat)
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 05-Nov-1999
 C; Accession: A60021
R; Wieczorek, D.F.; Hughes, S.R.
Brain Res. Mol. Brain Res. 10, 33-41, 1991
A; Title: Developmentally regulated cDNA expressed exclusively in neural tissue.
A; Reference number: A60021; MUID: 91278684; PMID: 1647480
A; Accession: A60021
A; Molecule type: mRNA
A; Residues: 1-267 <WIE>
A;Cross-references: EMBL:X52817; NID:g456549; PIDN:CAA37001.1; PID:g456550
C; Comment: This neuronal-specific mRNA was identified by hybridization to an
alpha-tropomyosin probe but does not show homology in amino acid sequence.
  Query Match
                         49.4%; Score 348; DB 2; Length 267;
  Best Local Similarity
                        64.0%; Pred. No. 8.9e-28;
  Matches
           64; Conservative
                             18; Mismatches
                                              18; Indels
                                                             0; Gaps
                                                                        0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
                   1 MDCVWSNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATIS 60
Db
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
Qу
             11111 1:11: 1:111111:111 1: :1:1 :111::
Db
          61 FRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 100
RESULT 3
A46583
neuroendocrine-specific protein, splice form A - human
N; Contains: neuroendocrine-specific protein, splice form B
C; Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999
C; Accession: A46583; I60903
R; Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers,
F.C.; Van de Ven, W.J.
J. Biol. Chem. 268, 13439-13447, 1993
```

A; Cross-references: GDB:203968; OMIM:600865

A; Map position: 14q21-14q22

```
A; Title: Cloning and expression of alternative transcripts of a novel
 neuroendocrine-specific gene and identification of its 135-kDa translational
 product.
 A; Reference number: A46583; MUID: 93293865; PMID: 7685762
 A; Accession: A46583
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: mRNA
 A; Residues: 1-776 < ROE1>
 A;Cross-references: GB:L10333; NID:g307306; PIDN:AAA59950.1; PID:g307307
 A; Accession: I60903
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: mRNA
A; Residues: 421-776 < ROE2>
A; Cross-references: GB:L10334; NID:g307308; PIDN:AAA59951.1; PID:g307309
 C; Genetics:
A; Gene: GDB: RTN1; NSP
A; Cross-references: GDB:203968; OMIM:600865
A; Map position: 14q21-14q22
  Query Match
                          47.8%; Score 337; DB 2; Length 776;
  Best Local Similarity
                          67.4%; Pred. No. 3.7e-26;
           62; Conservative 16; Mismatches
  Matches
                                                 14; Indels
                                                               0; Gaps
                                                                           0;
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
              586 KQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 645
Db
Qγ
          106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
              Db
          646 QAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 677
RESULT 4
T26216
hypothetical protein W06A7.3c - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C; Accession: T26216
R; Ainscough, R.
submitted to the EMBL Data Library, August 1996
A; Reference number: Z20173
A; Accession: T26216
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-2484 <WIL>
A;Cross-references: EMBL:Z78066; PIDN:CAB51467.1; GSPDB:GN00023; CESP:W06A7.3c
A; Experimental source: clone W06A7
C; Genetics:
A; Gene: CESP: W06A7.3c
A; Map position: 5
A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2
  Query Match
                         23.0%; Score 162; DB 2; Length 2484;
 Best Local Similarity 31.5%; Pred. No. 7.9e-08;
 Matches 34; Conservative 22; Mismatches
                                               48; Indels
                                                              4; Gaps
                                                                          1;
Qу
```

```
1:1:
                               1
                                      1:1::111 | | : : : : | : | : | : : : : |
          2269 HSILKHHGDAWIDF----KTVPPCVLDVIYWRDAKKSAIVLSLALLVLFVLAKYPLLTVV 2324
 Db
            85 AYIALALLSVTISFRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELV 132
 Qу
                1 1 1
                          2325 TYSLLLALGAAAGFRVFKKVEAQIKKTDSEHPFSEILAQDLTLPQEKV 2372
 Db
 RESULT 5
 T26215
 hypothetical protein W06A7.3a - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C; Accession: T26215
 R; Ainscough, R.
 submitted to the EMBL Data Library, August 1996
 A; Reference number: Z20173
 A; Accession: T26215
A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-2607 <WIL>
A;Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a
A; Experimental source: clone W06A7
 C; Genetics:
A; Gene: CESP: W06A7.3a
A; Map position: 5
A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2;
2586/2
  Query Match
                         22.6%; Score 159; DB 2; Length 2607;
  Best Local Similarity
                         34.5%; Pred. No. 1.7e-07;
  Matches
          30; Conservative 19; Mismatches 38; Indels
                                                              0; Gaps
                                                                           0:
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
              2409 KKEVLDVIYWRDAKKSAIVLSLALLVLFVLAKYPLLTVVTYSLLLALGAAAGFRVFKKVE 2468
Db
Qу
          106 QAIAKSDEGHPFRAYLESEVAISEELV 132
               1 1:1 111 | :::::|
Db
         2469 AQIKKTDSEHPFSEILAQDLTLPQEKV 2495
RESULT 6
T26213
hypothetical protein W06A7.3b - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C; Accession: T26213
R; Ainscough, R.
submitted to the EMBL Data Library, August 1996
A; Reference number: Z20173
A; Accession: T26213
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-222 <WIL>
A; Cross-references: EMBL: Z78066; PIDN: CAB01523.1; GSPDB: GN00023; CESP: W06A7.3b
A; Experimental source: clone W06A7
```

```
C; Genetics:
 A; Gene: CESP: W06A7.3b
 A; Map position: 5
 A; Introns: 27/1; 77/2; 201/2
   Query Match
                          22.4%; Score 158; DB 2; Length 222;
   Best Local Similarity 34.1%; Pred. No. 1.4e-08;
           29; Conservative 19; Mismatches 37; Indels
                                                                0; Gaps
                                                                            0;
           48 KVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQA 107
 Qy
              26 KILDVIYWRDAKKSAIVLSLALLVLFVLAKYPLLTVVTYSLLLALGAAAGFRVFKKVEAQ 85
 Db
 Qу
          108 IAKSDEGHPFRAYLESEVAISEELV 132
              | | : | | | | | | : : : : | |
Db
           86 IKKTDSEHPFSEILAQDLTLPQEKV 110
RESULT 7
S59439
probable membrane protein YDR233c - yeast (Saccharomyces cerevisiae)
N; Alternate names: hypothetical protein YD9934.17c
C; Species: Saccharomyces cerevisiae
C;Date: 30-Nov-1995 #sequence_revision 16-Feb-1996 #text_change 19-Apr-2002
C; Accession: S59439
R; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, March 1995
A; Reference number: S59423
A; Accession: S59439
A; Molecule type: DNA
A; Residues: 1-295 < MUR>
A;Cross-references: EMBL:Z48612; NID:g728671; PID:g728688; GSPDB:GN00004;
MIPS:YDR233c
A; Experimental source: strain AB972
C; Genetics:
A; Gene: MIPS: YDR233c
A; Cross-references: SGD:S0002641
A; Map position: 4R
C; Keywords: transmembrane protein
F;40-56/Domain: transmembrane #status predicted <TM1>
F;146-162/Domain: transmembrane #status predicted <TM2>
  Query Match
                         11.8%; Score 83.5; DB 2; Length 295;
 Best Local Similarity 26.3%; Pred. No. 0.73;
 Matches
          25; Conservative 17; Mismatches 32; Indels 21; Gaps
          41 QKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRI 100
Qy
                      111 11: :11 11 11 11 1 ::::
                                                          : 1 : 1
          12 QQQQQKSCNCDLLLWRNPVQTGKYFGGSLLALLILKKVNLITFFLKVAYTILFTT---- 66
Db
         101 YKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKY 135
Qу
               1 |: ::|
                                    ::: |: ||
          67 --GSIEFVSK-----LFLGQGLITKY 85
Db
```

RESULT 8 C88188

```
protein C18H9.5 [imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C; Accession: C88188
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode C. elegans: a platform for
 investigating biology.
 A; Reference number: A75000; MUID: 99069613; PMID: 9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and
 www_sanger.ac.uk/Projects/C_elegans/ for a list of authors
 A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103,
 1999; and Science 285, 1493, 1999
 A; Accession: C88188
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-464 <STO>
 A;Cross-references: GB:chr_II; PID:g722384; GSPDB:GN00020; CESP:C18H9.5
 C; Genetics:
 A; Gene: C18H9.5
 A; Map position: 2
   Query Match
                           11.6%; Score 81.5; DB 2; Length 464;
   Best Local Similarity
                           25.9%; Pred. No. 1.9;
            38; Conservative 28; Mismatches 50; Indels
  Matches
                                                                 31; Gaps
                                                                               8;
            8 VSCLRENFAVYSVSVGMHNLLLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKTGVVFG- 66
 Qу
                    1 :::
                                | :::| :|:
                                                : || :| |:
           41 ITCTNANMILMNFTVICMNDVIIEQKSF----SNQTHWLEKSSDISLTFSAAAVGAIFGT 96
 Db
           67 ASLFLLLS------ITVFSIVSV--TAYIALA----LLSVTISFRIYKGV------I 105
Qy
                             |||: ::|
                                       1 : 11
                                                   |: | |: |:::||
           97 VPAVTLISKYGIRKVLTVYGLLSAGGTLLMPLAVNYGLIPVLIA-RLFQGVGASILYSSI 155
Db
          106 QAIAKS----DEGHPFRAYLESEVAIS 128
Qу
                |::|
                            1 1:1 1
                       : |
Db
          156 GTISESWSPINEIGTFVAFLSSAFQIS 182
RESULT 9
T24228
hypothetical protein R166.2 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C; Accession: T24228
R; Matthews, P.
submitted to the EMBL Data Library, August 1995
A; Reference number: Z19859
A; Accession: T24228
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-618 <WIL>
A;Cross-references: EMBL:Z50795; PIDN:CAA90663.1; GSPDB:GN00020; CESP:R166.2
A; Experimental source: clone R166
C; Genetics:
A; Gene: CESP:R166.2
A; Map position: 2
```

```
A; Introns: 41/3; 86/3; 118/3; 164/3; 207/3; 554/3
  Query Match
                       11.2%; Score 79; DB 2; Length 618;
  Best Local Similarity 29.1%; Pred. No. 4.7;
  Matches 32; Conservative 22; Mismatches
                                            32; Indels 24; Gaps
                                                                    6;
Qу
          51 DLLYWRDIKK-TGVVFGASLF-LLLSLTVF------SIVSVTAYIALAL----- 91
             Db
         353 DIQFWNNRKDLVGLSVRSVLFNIFQSLIVFLYICDNETNTMVKVTVGIGLLIECWKIPKV 412
Qу
          92 LSVTISFR-IYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
            Db
         413 MNVSIDWQNKWFGVIPRLVISDKG----SYVESETKIYDQMAFKYLGWAL 458
RESULT 10
A72258
hypothetical protein - Thermotoga maritima (strain MSB8)
C; Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C; Accession: A72258
R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft,
D.H.; Hickey, E.K.; Peterson, J.D.; Nelson, W.C.; Ketchum, K.A.; McDonald, L.;
Utterback, T.R.; Malek, J.A.; Linher, K.D.; Garrett, M.M.; Stewart, A.M.;
Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; Heidelberg, J.;
Sutton, G.G.; Fleischmann, R.D.; White, O.; Salzberg, S.L.; Smith, H.O.; Venter,
J.C.; Fraser, C.M.
Nature 399, 323-329, 1999
A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.
A; Reference number: A72200; MUID: 99287316; PMID: 10360571
A; Accession: A72258
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-458 <ARN>
A;Cross-references: GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD36479.1;
PID:g4981972; TIGR:TM1408
A; Experimental source: strain MSB8
C; Genetics:
A; Gene: TM1408
 Query Match
                      11.1%; Score 78.5; DB 2; Length 458;
 Best Local Similarity 26.6%; Pred. No. 3.8;
 Matches 42; Conservative 17; Mismatches 34; Indels 65; Gaps
                                                                    9;
Qу
         14 NFAVY----SVSVGMHNLLLLEGRSWQEMDGOKKH-----WKD 47
            || :
                  196 NFLILSYLRSSIRIGFDFLLL----TRKHPQLLFIGYFYYLSIWIDNFIAWKV 244
Db
Qу
         48 KVVDLL-----YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFR 99
            Db
        245 KGIEIAPGFFMSPEY--DIPK----FMASLFFIPSLVVFN-----LSMETVFQR 287
        100 IYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
Qу
             |||::|:| ||: || ||: :| || :||
Db
        288 NYKGLMQSIV-SDK--PMRVISENLKKLSLSLRHAFSN 322
```

```
RESULT 11
F64924
probable thiosulfate-dithiol sulfurtransferase (EC 2.8.1.5) - Escherichia coli
(strain K-12)
C; Species: Escherichia coli
C; Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 03-Jun-2002
C; Accession: F64924
R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;
Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor,
J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID: 97426617; PMID: 9278503
A; Accession: F64924
A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-261 <BLAT>
A; Cross-references: GB: AE000262; GB: U00096; NID: g1787955; PIDN: AAC74740.1;
PID:g1787959; UWGP:b1670
A; Experimental source: strain K-12, substrain MG1655
C; Superfamily: hyaC protein
C; Keywords: sulfurtransferase; transmembrane protein
F;30-46/Domain: transmembrane #status predicted <TM01>
F;83-99/Domain: transmembrane #status predicted <TM02>
F;112-128/Domain: transmembrane #status predicted <TM03>
F;187-203/Domain: transmembrane #status predicted <TM04>
F;224-240/Domain: transmembrane #status predicted <TM05>
                          11.0%; Score 77.5; DB 2; Length 261;
  Query Match
                         27.6%; Pred. No. 2.6;
  Best Local Similarity
           29; Conservative 16; Mismatches
                                                  43; Indels
                                                                17; Gaps
  Matches
           22 VGMHNLLLLEGRSWQEMD-GQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLL----SLT 76
Qу
                          | ||:||
              :1:1 11
                      44 LGLHALLRARGVKKSATDHGEKIYLYSKAVRLWHWSN-----ALLFVLLLASGLIN 94
Db
           77 VFSIVSVTAYIALALLSVTISFRI---YKGVIQAIAKSDEGHPFR 118
Qу
               1::| || :| :
                                 1:
                                       : | : | | | || :|
           95 HFAMVGATAVKSLVAVHEVCGFLLLACWLGFVLINAVGDNGHHYR 139
Db
RESULT 12
T49359
hypothetical protein B1D1.130 [imported] - Neurospora crassa
C; Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence revision 02-Jun-2000 #text change 02-Jun-2000
C:Accession: T49359
R; Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.;
Nyakatura, G.; Mewes, H.W.; Mannhaupt, G.
submitted to the Protein Sequence Database, May 2000
A; Reference number: Z25022
A; Accession: T49359
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-583 <SCH>
A;Cross-references: EMBL:AL355927; GSPDB:GN00116; NCSP:B1D1.130
```

```
A; Experimental source: BAC clone B1D1; strain OR74A
C; Genetics:
A; Gene: NCSP: B1D1.130
A; Map position: 6
A; Introns: 44/1
                         11.0%; Score 77.5; DB 2; Length 583;
  Query Match
  Best Local Similarity 27.5%; Pred. No. 6.2;
          28; Conservative 13; Mismatches
                                                 28; Indels
           52 LLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKS 111
Qу
              Db
           67 LCYW-PLERRLLVLHALLLLLLSLEHYS-----AYTRVLLLHITSSLNL------ 109
          112 DEGHPFRAYLESEVAIS-----EELVQKYSNSALG 141
QУ
                 | | :: || ::
                                          110 ----PLRVLVDDEVRVAKAIAWMAKDINPEELIQKRIEECAG 147
Db
RESULT 13
C95920
hypothetical membrane protein [imported] - Sinorhizobium meliloti (strain 1021)
magaplasmid pSymB
C; Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence revision 24-Aug-2001 #text change 30-Sep-2001
C; Accession: C95920
R; Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter,
F.J.; Hernandez-Lucas, I.; Becker, A.; Cowie, A.; Gouzy, J.; Golding, B.;
Puhler, A.
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A; Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.
A; Reference number: A95842; MUID:21396508; PMID:11481431
A; Accession: C95920
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-481 < KUR>
A; Cross-references: GB: AL591985; PIDN: CAC49027.1; PID: q15140512; GSPDB: GN00167
A; Experimental source: strain 1021, megaplasmid pSymB
R; Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.;
Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; Bothe, G.; Boutry,
M.; Bowser, L.; Buhrmester, J.; Cadieu, E.; Capela, D.; Chain, P.; Cowie, A.;
Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Gloux, S.; Godrie, T.;
Goffeau, A.; Golding, B.; Gouzy, J.; Gurjal, M.; Hernandez-Lucas, I.; Hong, A.;
Huizar, L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A; Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.;
Lelaure, V.; Masuy, D.; Palm, C.; Peck, M.C.; Pohl, T.M.; Portetelle, D.;
Purnelle, B.; Ramsperger, U.; Surzycki, R.; Thebault, P.; Vandenbol, M.;
Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.; Batut, J.
A; Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A; Reference number: A96039; MUID: 21368234; PMID: 11474104
A; Contents: annotation
C; Genetics:
A; Gene: SMb21048
A; Genome: plasmid
```

```
Query Match
                         10.9%; Score 76.5; DB 2; Length 481;
  Best Local Similarity 37.7%; Pred. No. 6.4;
           23; Conservative
                                8; Mismatches
                                                27; Indels
                                                               3; Gaps
                                                                           2;
          53 LYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD 112
Qу
                        Db
          238 VYWRKTKSR--VAQALAFLLLVLLLLSTSSV-AYVGLAVLSIPVALSISWSFLSGRMDKD 294
Qy
         113 E 113
Db
         295 E 295
RESULT 14
F64665
glucose-6-phosphate isomerase (EC 5.3.1.9) - Helicobacter pylori (strain 26695)
C; Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence revision 09-Aug-1997 #text change 16-Jul-1999
C; Accession: F64665
R; Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;
Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;
Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;
Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.; Fitzegerald,
L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback,
T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.;
Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A; Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.;
Fraser, C.M.; Venter, J.C.
A; Title: The complete genome sequence of the gastric pathogen Helicobacter
pylori.
A; Reference number: A64520; MUID: 97394467; PMID: 9252185
A; Accession: F64665
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-545 < TOM>
A;Cross-references: GB:AE000622; GB:AE000511; NID:g2314317; PIDN:AAD08211.1;
PID:q2314323; TIGR:HP1166
C; Superfamily: glucose-6-phosphate isomerase
C; Keywords: intramolecular oxidoreductase; isomerase
 Query Match
                         10.7%; Score 75.5; DB 2; Length 545;
 Best Local Similarity
                        25.3%; Pred. No. 9.2;
           37; Conservative 20; Mismatches
                                                44; Indels
Qy
          23 GMHNLL-----LLEGRSWQEMDGQ---KKHWKDKVVDLLYWRDIKKTGVVFGASL 69
                             ::|:|::| |: | ||: || : |
                                                                Db
         411 GHHEILFSNVLAQAQAFMKGKSYEEALGELLFKGLDKDEAKDLAHHR-----VFFGNRP 464
          70 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD----- 112
Qу
                    | :: | :||
                                  | :
Db
         465 SNILLLEKISPSNIGALVALYEHKVFV-----QGVIWDINSFDQWGVELGKELAVPILQE 519
         113 -EGHPFRAYLESEVAISEELVQKYSN 137
Qу
                  || :|
              :: |:: | |
Db
         520 LEGHKSNAYFDSS---TKHLIELYKN 542
```

```
RESULT 15
E71851
glucose-6-phosphate isomerase - Helicobacter pylori (strain J99)
C; Species: Helicobacter pylori
A; Variety: strain J99
C;Date: 12-Feb-1999 #sequence revision 12-Feb-1999 #text change 16-Jul-1999
C; Accession: E71851
R; Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.;
Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.;
Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.;
Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.
Nature 397, 176-180, 1999
A; Title: Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.
A; Reference number: A71800; MUID: 99120557; PMID: 9923682
A; Accession: E71851
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-545 <ARN>
A; Cross-references: GB: AE001536; GB: AE001439; NID: g4155675; PIDN: AAD06664.1;
PID:g4155679
A; Experimental source: strain J99
C; Genetics:
A; Gene: pgi
C; Superfamily: glucose-6-phosphate isomerase
  Query Match
                          10.6%; Score 74.5; DB 2; Length 545;
  Best Local Similarity
                         25.3%; Pred. No. 12;
  Matches
           37; Conservative 19; Mismatches
                                                  45; Indels
                                                               45; Gaps
           23 GMHNLL-----LLEGRSWQEMDGQ---KKHWKDKVVDLLYWRDIKKTGVVFGASL 69
Qу
              1 | : |
                             ::|:|::| |: | ||: ||: |
                                                                  1 11
Db
          411 GHHEILFSNVLAQAQAFMKGKSYEEALGELLSKGLDKDEAKDLAHHR-----VFFGNRP 464
Qу
           70 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD----- 112
                : | | : | : | |
                                    | :
                                             : | | | |
Db
          465 SNILLLEKISPSNIGALVALYEHKVFV----QGVIWDINSFDQWGVELGKELAVPILQE 519
          113 -EGHPFRAYLESEVAISEELVQKYSN 137
Qу
                   || :|
                            : |:: | |
Db
          520 LEGHKSNAYFDSS---TRHLIELYKN 542
Search completed: September 3, 2004, 16:10:04
```

Job time : 8.01088 secs

## GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

September 3, 2004, 16:09:02; Search time 13.9182 Seconds Run on:

(without alignments)

3191.803 Million cell updates/sec

Title: US-09-830-972-32

Perfect score: 705

Sequence: 1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVOKYSNSALG 141

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 1298764

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

/cgn2 6/ptodata/1/pubpaa/US07 PUBCOMB.pep:\*

/cgn2 6/ptodata/1/pubpaa/PCT NEW PUB.pep:\*

/cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pep:\*

4: /cgn2 6/ptodata/1/pubpaa/US06 PUBCOMB.pep:\*

/cgn2 6/ptodata/1/pubpaa/US07 NEW PUB.pep:\*

6:

/cgn2 6/ptodata/1/pubpaa/PCTUS PUBCOMB.pep:\* 7: /cgn2 6/ptodata/1/pubpaa/US08 NEW PUB.pep:\*

/cgn2 6/ptodata/1/pubpaa/US08 PUBCOMB.pep:\*

/cgn2 6/ptodata/1/pubpaa/US09A PUBCOMB.pep:\* 9:

10: /cgn2 6/ptodata/1/pubpaa/US09B PUBCOMB.pep:\*

11: /cgn2 6/ptodata/1/pubpaa/US09C PUBCOMB.pep:\*

12: /cgn2 6/ptodata/1/pubpaa/US09 NEW PUB.pep:\*

13: /cgn2 6/ptodata/1/pubpaa/US10A PUBCOMB.pep:\*

14: /cgn2 6/ptodata/1/pubpaa/US10B PUBCOMB.pep:\*

15: /cgn2 6/ptodata/1/pubpaa/US10C PUBCOMB.pep:\*

16: /cgn2 6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*

/cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\* 17:

18: /cgn2 6/ptodata/1/pubpaa/US60 PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ક

Result Query

> No. Score Match Length DB ID

Description

1	510	72.3	199	9	US-09-893-348-21	Sequence 21, Appl
2	503	71.3	118	15		Sequence 1568, Ap
3	503	71.3	199	9	US-09-893-348-25	Sequence 25, Appl
4	503	71.3	199	12	US-10-660-946-1	Sequence 1, Appli
5	503	71.3	199	12	US-09-978-360A-467	Sequence 467, App
6	503	71.3	199	16		Sequence 11, Appl
7	496	70.4	199	12		Sequence 9, Appli
8	447	63.4	1163	9	US-09-893-348-18	Sequence 18, Appl
9	447	63.4	1192	9	US-09-789-386-2	Sequence 2, Appli
10	447	63.4	1192	9	US-09-758-140-6	Sequence 6, Appli
11	447	63.4	1192	9	US-09-893-348-23	Sequence 23, Appl
12	447	63.4	1192	9	US-09-972-599A-6	Sequence 6, Appli
13	447	63.4	1192	12	US-10-267-502-429	Sequence 429, App
14	447	63.4	1192	14	US-10-060-036-71	Sequence 71, Appl
15	447	63.4	1192	16	US-10-327-213-9	Sequence 9, Appli
16	447	63.4	1192	16	US-10-466-258-9	Sequence 9, Appli
17	443	62.8	360	9	US-09-893-348-20	Sequence 20, Appl
18	443	62.8	373	9	US-09-789-386-6	Sequence 6, Appli
19	443	62.8	373	9	US-09-765-205-6	Sequence 6, Appli
20	443	62.8	373	9	US-09-893-348-24	Sequence 24, Appl
21	443	62.8	373	12	US-10-408-967-8	Sequence 8, Appli
22	443	62.8	373	14	US-10-060-036-72	Sequence 72, Appl
23	443	62.8	373	16	US-10-466-258-4	Sequence 4, Appli
24	443	62.8	379	14	US-10-205-194-164	Sequence 164, App
25	440	62.4	1192	12	US-10-408-967-7	Sequence 7, Appli
26	436.5	61.9	1163	12	US-10-267-502-431	Sequence 431, App
27	349	49.5	208	12	US-10-660-946-7	Sequence 7, Appli
28	348	49.4	267	12	US-10-660 <b>-</b> 946-8	Sequence 8, Appli
29	348	49.4	267	14	US-10-205-194-127	Sequence 127, App
30	337	47.8	356	12	US-10-660-946-6	Sequence 6, Appli
31	337	47.8	776	12	US-10-660-946-5	Sequence 5, Appli
32	337	47.8	776	12	US-10-267-502-430	Sequence 430, App
33	337	47.8	777	14	US-10-205-219-93	Sequence 93, Appl
34	337	47.8	780	12	US-10-267-502-432	Sequence 432, App
35	309	43.8	236	9	US-09-729-674-20	Sequence 20, Appl
36	309	43.8	236	9	US-09-765-205-26	Sequence 26, Appl
37	309	43.8	236	12	US-10-408-967-2	Sequence 2, Appli
38	309	43.8	266	12	US-10-276-774-2330	Sequence 2330, Ap
39	309	43.8	269	14	US-10-106-698-6222	Sequence 6222, Ap
40	305	43.3	241	12	US-10-660-946-3	Sequence 3, Appli
41	304	43.1	593	15	US-10-108-260A-2892	Sequence 2892, Ap
42	269	38.2	161	9	US-09-925-302-808	Sequence 808, App
43	269	38.2	161	12	US-09-925-302-808	Sequence 808, App
44	259.5	36.8	204	12	US-10-424-599-194160	Sequence 194160,
45	256	36.3	234	12	US-10-424-599-200840	Sequence 200840,

# ALIGNMENTS

## RESULT 1

US-09-893-348-21

- ; Sequence 21, Application US/09893348
  ; Patent No. US20020072493A1
- ; GENERAL INFORMATION:
- ; APPLICANT: EISENBACH-SCHWARTZ, Michal

```
; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
  APPLICANT: MOALEM, Gila
  TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
  CURRENT FILING DATE: 2001-06-28
  PRIOR APPLICATION NUMBER: US 09/314,161
  PRIOR FILING DATE: 1999-05-19
  PRIOR APPLICATION NUMBER: US 09/218,277
  PRIOR FILING DATE: 1998-12-22
  PRIOR APPLICATION NUMBER: PCT/US98/14715
  PRIOR FILING DATE: 1998-07-21
  PRIOR APPLICATION NUMBER: IL 124500
  PRIOR FILING DATE: 1998-05-19
  NUMBER OF SEQ ID NOS: 29
  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
   LENGTH: 199
   TYPE: PRT
   ORGANISM: Rattus norvegicus
US-09-893-348-21
  Query Match
                        72.3%; Score 510; DB 9; Length 199;
 Best Local Similarity 99.0%; Pred. No. 5.7e-51;
 Matches 103; Conservative 0; Mismatches 1; Indels
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
             Db
           1 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 2
US-10-264-237-1568
; Sequence 1568, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
  FILE REFERENCE: PA131P1
  CURRENT APPLICATION NUMBER: US/10/264,237
  CURRENT FILING DATE: 2002-10-04
  PRIOR APPLICATION NUMBER: PCT/US01/16450
  PRIOR FILING DATE: 2001-05-18
  PRIOR APPLICATION NUMBER: US 60/205,515
  PRIOR FILING DATE: 2000-05-19
  NUMBER OF SEQ ID NOS: 2876
  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1568
   LENGTH: 118
   TYPE: PRT
```

```
ORGANISM: Homo sapiens
   FEATURE:
   NAME/KEY: MISC FEATURE
   LOCATION: (118)
   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino
acids
US-10-264-237-1568
  Query Match
                        71.3%; Score 503; DB 15; Length 118;
  Best Local Similarity 98.1%; Pred. No. 1.9e-50;
 Matches 102; Conservative 1; Mismatches
                                              1; Indels 0; Gaps
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
             Db
           1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 3
US-09-893-348-25
; Sequence 25, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
 TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
  PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
  PRIOR APPLICATION NUMBER: US 09/218,277
  PRIOR FILING DATE: 1998-12-22
  PRIOR APPLICATION NUMBER: PCT/US98/14715
  PRIOR FILING DATE: 1998-07-21
  PRIOR APPLICATION NUMBER: IL 124500
  PRIOR FILING DATE: 1998-05-19
 NUMBER OF SEQ ID NOS: 29
  SOFTWARE: PatentIn version 3.1
; SEO ID NO 25
   LENGTH: 199
   TYPE: PRT
   ORGANISM: Homo sapiens
US-09-893-348-25
 Query Match
                        71.3%; Score 503; DB 9; Length 199;
 Best Local Similarity 98.1%; Pred. No. 3.7e-50;
 Matches 102; Conservative 1; Mismatches
                                                            0; Gaps
                                                                       0;
                                             1; Indels
```

Qy

```
Db
           1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Qу
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 4
US-10-660-946-1
; Sequence 1, Application US/10660946
; Publication No. US20040063131A1
   GENERAL INFORMATION:
        APPLICANT: Bandman, Olga
                  Au-Young, Janice
                  Goli, Surya K.
                  Hillman, Jennifer L.
        TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
        NUMBER OF SEQUENCES: 9
        CORRESPONDENCE ADDRESS:
            ADDRESSEE: Incyte Pharmaceuticals, Inc.
             STREET: 3174 Porter Drive
            CITY: Palo Alto
             STATE: CA
             COUNTRY: U.S.
             ZIP: 94304
     COMPUTER READABLE FORM:
            MEDIUM TYPE: Diskette
             COMPUTER: IBM Compatible
            OPERATING SYSTEM: DOS
            SOFTWARE: FastSEQ Version 1.5
      CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/10/660,946
            FILING DATE: 12-Sep-2003
       PRIOR APPLICATION DATA:
            APPLICATION NUMBER: US/09/228,213A
             FILING DATE: <Unknown>
            APPLICATION NUMBER: 08/700,607
             FILING DATE: <Unknown>
       ATTORNEY/AGENT INFORMATION:
            NAME: Billings, Lucy J.
            REGISTRATION NUMBER: 36,749
            REFERENCE/DOCKET NUMBER: PF-0114 US
        TELECOMMUNICATION INFORMATION:
            TELEPHONE: 415-855-0555
            TELEFAX: 415-845-4166
   INFORMATION FOR SEQ ID NO: 1:
        SEQUENCE CHARACTERISTICS:
            LENGTH: 199 amino acids
            TYPE: amino acid
            STRANDEDNESS: single
            TOPOLOGY: linear
       MOLECULE TYPE: peptide
        IMMEDIATE SOURCE:
            LIBRARY: <Unknown>
            CLONE: Consensus
        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
```

```
Query Match
                        71.3%; Score 503; DB 12; Length 199;
  Best Local Similarity 98.1%; Pred. No. 3.7e-50;
  Matches 102; Conservative 1; Mismatches
                                              1; Indels
                                                             0; Gaps
                                                                        0;
QУ
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
             1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 5
US-09-978-360A-467
; Sequence 467, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouqueleret, Lydie
 APPLICANT: Jobert, Severin
  APPLICANT: Clusel, Catherine
  TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal
Peptides
  FILE REFERENCE: 56.US4.CIP
  CURRENT APPLICATION NUMBER: US/09/978,360A
  CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
  PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
  PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
  PRIOR FILING DATE: 1998-04-13
  PRIOR APPLICATION NUMBER: US 60/096,116
  PRIOR FILING DATE: 1998-08-10
  PRIOR APPLICATION NUMBER: US 60/099,273
  PRIOR FILING DATE: -09-04
  PRIOR APPLICATION NUMBER: US 09/191,997
  PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
  PRIOR FILING DATE: 1998-12-17
  PRIOR APPLICATION NUMBER: PCT/IB98/02122
  PRIOR FILING DATE: 1998-12-17
  PRIOR APPLICATION NUMBER: US 09/247,155
  PRIOR FILING DATE: 1999-02-09
  Remaining Prior Application data removed - See File Wrapper or PALM.
  NUMBER OF SEQ ID NOS: 810
 SOFTWARE: Patent.pm
; SEQ ID NO 467
   LENGTH: 199
   TYPE: PRT
   ORGANISM: Homo sapiens
```

```
FEATURE:
   NAME/KEY: SIGNAL
   LOCATION: -42..-1
US-09-978-360A-467
 Query Match
                       71.3%; Score 503; DB 12; Length 199;
 Best Local Similarity
                       98.1%; Pred. No. 3.7e-50;
 Matches 102; Conservative 1; Mismatches
                                            1; Indels
                                                         0; Gaps
                                                                   0:
Qу
         38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
            1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Db
Qу
         98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
            Db
         61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 6
US-10-466-258-11
; Sequence 11, Application US/10466258
; Publication No. US20040132096A1
; GENERAL INFORMATION:
  APPLICANT: GLAXO GROUP LIMITED
  TITLE OF INVENTION: ASSAY
  FILE REFERENCE: P80966 GCW
  CURRENT APPLICATION NUMBER: US/10/466,258
  CURRENT FILING DATE: 2003-07-15
  NUMBER OF SEQ ID NOS: 13
  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
   LENGTH: 199
   TYPE: PRT
   ORGANISM: Homo sapiens
US-10-466-258-11
 Query Match
                      71.3%; Score 503; DB 16; Length 199;
 Best Local Similarity
                      98.1%; Pred. No. 3.7e-50;
 Matches 102; Conservative
                          1; Mismatches
                                           1; Indels
                                                        0; Gaps
Qу
         38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
            Db
          1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
         98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
            61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVOKYSNSALG 104
Db
RESULT 7
US-10-408-967-9
; Sequence 9, Application US/10408967
; Publication No. US20040063161A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Yan, Riqiang
; APPLICANT: Lu, Yifeng
```

```
; TITLE OF INVENTION: Compositions and Methods of Treating Alzheimer's Disease
  FILE REFERENCE: 00925
  CURRENT APPLICATION NUMBER: US/10/408,967
  CURRENT FILING DATE: 2003-04-08
  NUMBER OF SEQ ID NOS: 9
  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
   LENGTH: 199
   TYPE: PRT
   ORGANISM: Homo sapiens
US-10-408-967-9
  Query Match
                        70.4%; Score 496; DB 12; Length 199;
  Best Local Similarity 97.1%; Pred. No. 2.4e-49;
 Matches 101; Conservative 1; Mismatches
                                                             0; Gaps
                                               2; Indels
                                                                        0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
             1 MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Db
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
QУ
             Db
          61 FRIYKGVIQAIQKSDEGHPFPAYLESEVAISEELVQKYSNSALG 104
RESULT 8
US-09-893-348-18
; Sequence 18, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
  APPLICANT: EISENBACH-SCHWARTZ, Michal
 APPLICANT: COHEN, Irun R.
  APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
  CURRENT APPLICATION NUMBER: US/09/893,348
  CURRENT FILING DATE: 2001-06-28
  PRIOR APPLICATION NUMBER: US 09/314,161
  PRIOR FILING DATE: 1999-05-19
  PRIOR APPLICATION NUMBER: US 09/218,277
  PRIOR FILING DATE: 1998-12-22
  PRIOR APPLICATION NUMBER: PCT/US98/14715
  PRIOR FILING DATE: 1998-07-21
  PRIOR APPLICATION NUMBER: IL 124500
  PRIOR FILING DATE: 1998-05-19
  NUMBER OF SEQ ID NOS: 29
  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
   LENGTH: 1163
   TYPE: PRT
   ORGANISM: Rattus norvegicus
US-09-893-348-18
```

```
Best Local Similarity 96.9%; Pred. No. 1.3e-42;
         93; Conservative 0; Mismatches 3; Indels
                                                          0; Gaps
                                                                    0;
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
            Db
         973 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1032
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qy
            1033 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1068
Db
RESULT 9
US-09-789-386-2
; Sequence 2, Application US/09789386
; Patent No. US20020010324A1
; GENERAL INFORMATION:
; APPLICANT: MICHALOVICH, DAVID
; APPLICANT: PRINJHA, RABINDER KUMAR
; TITLE OF INVENTION: NOVEL COMPOUNDS
  FILE REFERENCE: GP-30165-C1
  CURRENT APPLICATION NUMBER: US/09/789,386
  CURRENT FILING DATE: 2001-02-21
  PRIOR APPLICATION NUMBER: U.K. 9916898.1
  PRIOR FILING DATE: 1999-07-19
  PRIOR APPLICATION NUMBER: U.K. 9816024.5
  PRIOR FILING DATE: 1998-07-22
  PRIOR APPLICATION NUMBER: US 09/359,208
  PRIOR FILING DATE: 1999-07-22
 NUMBER OF SEQ ID NOS: 6
  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
   LENGTH: 1192
   TYPE: PRT
   ORGANISM: HOMO SAPIENS
US-09-789-386-2
 Query Match
                       63.4%; Score 447; DB 9; Length 1192;
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
                            0; Mismatches
 Matches 93; Conservative
                                           3; Indels
                                                                    0;
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
QУ
              1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
Qу
        106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
            1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
Db
RESULT 10
US-09-758-140-6
; Sequence 6, Application US/09758140
; Patent No. US20020012965A1
; GENERAL INFORMATION:
; APPLICANT: Strittmatter, Stephen M.
```

```
; TITLE OF INVENTION: No. US20020012965Alo Receptor-Mediated Blockade of
; FILE REFERENCE: 44574-5073-US
; CURRENT APPLICATION NUMBER: US/09/758,140
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,707
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: US 60/207,366
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,378
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
  LENGTH: 1192
  TYPE: PRT
   ORGANISM: Homo sapiens
US-09-758-140-6
                       63.4%; Score 447; DB 9; Length 1192;
 Query Match
  Best Local Similarity 96.9%; Pred. No. 1.3e-42;
 Matches 93; Conservative 0; Mismatches 3; Indels 0; Gaps
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qy
             Db
        1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
RESULT 11
US-09-893-348-23
; Sequence 23, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: COHEN, Irun R. ; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
  PRIOR FILING DATE: 1998-07-21
  PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 23
  LENGTH: 1192
   TYPE: PRT
   ORGANISM: Homo sapiens
US-09-893-348-23
 Query Match
                      63.4%; Score 447; DB 9; Length 1192;
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
         93; Conservative 0; Mismatches
                                            3; Indels
                                                                   0 ;
Qy
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
            Db
       1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
        106 QAIAKSDEGHPFRAYLESEVAISEELVOKYSNSALG 141
Qу
            Dh
       1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
RESULT 12
US-09-972-599A-6
; Sequence 6, Application US/09972599A
; Patent No. US20020077295A1
; GENERAL INFORMATION:
; APPLICANT: STRITTMATTER, STEPHEN M.
  TITLE OF INVENTION: NOGO-RECEPTOR-MEDIATED BLOCKADE OF AXONAL GROWTH
; FILE REFERENCE: C077 CIP US
  CURRENT APPLICATION NUMBER: US/09/972,599A
  CURRENT FILING DATE: 2001-10-06
  PRIOR APPLICATION NUMBER: PCT/US01/01041
; PRIOR FILING DATE: 2001-01-12
  PRIOR APPLICATION NUMBER: 09/758,140
  PRIOR FILING DATE: 2001-01-12
  PRIOR APPLICATION NUMBER: 60/236,378
  PRIOR FILING DATE: 2000-09-29
  PRIOR APPLICATION NUMBER: 60/207,366
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: 60/175,707
  PRIOR FILING DATE: 2000-01-12
  NUMBER OF SEQ ID NOS: 57
  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
  LENGTH: 1192
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-972-599A-6
                      63.4%; Score 447; DB 9; Length 1192;
 Query Match
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
         93; Conservative 0; Mismatches
                                           3; Indels
Qу
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
            Db
       1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
QУ
        106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
```

```
RESULT 13
US-10-267-502-429
; Sequence 429, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
  TITLE OF INVENTION: Obesity Linked Genes
  FILE REFERENCE: LSD-07416
  CURRENT APPLICATION NUMBER: US/10/267,502
  CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 429
   LENGTH: 1192
   TYPE: PRT
   ORGANISM: Homo sapiens
US-10-267-502-429
 Query Match
                        63.4%; Score 447; DB 12; Length 1192;
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
 Matches 93; Conservative 0; Mismatches
                                              3; Indels
                                                             0; Gaps
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
                Db
        1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
             Db
        1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
RESULT 14
US-10-060-036-71
; Sequence 71, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
  APPLICANT: Benson, Darin R.
  APPLICANT: Kalos, Michael D.
  APPLICANT: Lodes, Michael J.
  APPLICANT: Persing, David H.
  APPLICANT: Hepler, William T.
  APPLICANT: Jiang, Yuqiu
  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
  TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
  FILE REFERENCE: 210121.566
  CURRENT APPLICATION NUMBER: US/10/060,036
  CURRENT FILING DATE: 2002-01-30
  NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
   LENGTH: 1192
   TYPE: PRT
   ORGANISM: Homo sapiens
```

```
Query Match
                     63.4%; Score 447; DB 14; Length 1192;
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
 Matches 93; Conservative 0; Mismatches 3; Indels
                                                        0; Gaps
                                                                  0;
Qγ
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
            Db
       1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
QУ
        106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
            Db
       1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
RESULT 15
US-10-327-213-9
; Sequence 9, Application US/10327213
; Publication No. US20040121341A1
; GENERAL INFORMATION:
; APPLICANT: FILBIN, MARIE T.
; APPLICANT: DOMENICONI, MARCO
; APPLICANT: CAO, ZIXUAN
; TITLE OF INVENTION: INHIBITORS OF MYELIN-ASSOCIATED GLYCOPROTEIN (MAG)
 TITLE OF INVENTION: ACTIVITY FOR REGULATING NEURAL GROWTH AND REGENERATION
  FILE REFERENCE: CUNY/003
 CURRENT APPLICATION NUMBER: US/10/327,213
 CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
  LENGTH: 1192
  TYPE: PRT
   ORGANISM: Homo sapiens
US-10-327-213-9
 Query Match
                      63.4%; Score 447; DB 16; Length 1192;
 Best Local Similarity 96.9%; Pred. No. 1.3e-42;
 Matches 93; Conservative 0; Mismatches
                                          3; Indels
                                                      0; Gaps
         46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
            Db
       1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
Qy
        106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
            Db
       1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
```

Search completed: September 3, 2004, 16:30:09 Job time: 14.9182 secs

## GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2004, 15:59:39; Search time 8.86221 Seconds

(without alignments)

5019.975 Million cell updates/sec

Title:

US-09-830-972-32

Perfect score: 705

Sequence: 1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVQKYSNSALG 141

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:\*

1: sp archea:\*

2: sp bacteria:\*

3: sp fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp mammal:\*

7: sp mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*
11: sp\_rodent:\*

12: sp\_virus:\*

13: sp vertebrate:\*

14: sp unclassified:\*

15: sp rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result Query

No. Score Match Length DB ID Description \_\_\_\_\_\_

1	460	65.2	199	13	Q7T224	Q7t224 gallus gall
2	447	63.4	578	11	Q80 <b>W</b> 95	Q80w95 mus musculu
3	447	63.4	639	11	Q8K290	Q8k290 mus musculu
4	447	63.4	986	4	Q8IUA4	Q8iua4 homo sapien
5	447	63.4	1046	11	Q8BGK7	Q8bgk7 mus musculu
6	447	63.4	1162	11	Q8BGM9	Q8bgm9 mus musculu
7	443	62.8	356	11	Q8BH78	Q8bh78 mus musculu
8	443	62.8	375	11	Q8BHF5	Q8bhf5 mus musculu
9	443	62.8	392	4	Q96B16	Q96b16 homo sapien
10	436.5	61.9	1163	11	Q8K3G8	Q8k3g8 mus musculu
11	432.5	61.3	357	11	Q8K3G7	Q8k3g7 mus musculu
12	417	59.1	184	6	Q7YRW9	Q7yrw9 bos taurus
13	386	54.8	179	6	Q9GM33	Q9gm33 macaca fasc
14	384	54.5	214	13	Q7T222	Q7t222 carassius a
15	348	49.4	199	4	Q9BQ59	Q9bq59 homo sapien
16	348	49.4	267	11	Q63765	Q63765 rattus sp.
17	337	47.8	780	11	Q8K4S4	Q8k4s4 mus musculu
18	337	47.8	780	11	Q8K0T0	Q8k0t0 mus musculu
19	325	46.1	208	13	Q90637	Q90637 gallus gall
20	320	45.4	760	13	Q90638	Q90638 gallus gall
21	308	43.7	236	11	Q8VBU0	Q8vbu0 rattus norv
22	308	43.7	237	11	Q8C6D5	Q8c6d5 mus musculu
23	308	43.7	643	11	Q8CCU2	Q8ccu2 mus musculu
24	283.5	40.2	221	13	Q7ZUD6	Q7zud6 brachydanio
25	255	36.2	234	5	Q9VMW3	Q9vmw3 drosophila
26	253	35.9	224	5	Q9VMW1	Q9vmw1 drosophila
27	252	35.7	222	5	Q9VMW4	Q9vmw4 drosophila
28	252	35.7	595	5	Q9VMV9	Q9vmv9 drosophila
29	250	35.5	202	5	Q9VMW2	Q9vmw2 drosophila
30	162	23.0	2484	5	Q9U347	Q9u347 caenorhabdi
31	159	22.6	2607	5	Q23187	Q23187 caenorhabdi
32	158	22.4	222	5	Q23188	Q23188 caenorhabdi
33	104	14.8	107	13	Q7T223	Q7t223 carassius a
34	102.5	14.5	154	5	Q9VIB7	Q9vib7 drosophila
35	102.5	14.5	158	5	Q24199	Q24199 drosophila
36	85	12.1	457	10	Q8LDS3	Q8lds3 arabidopsis
37	83.5	11.8	295	3	Q04947	Q04947 saccharomyc
38	82.5	11.7	1247	16	Q88QU8	Q88qu8 pseudomonas
39	81.5	11.6	464	5	Q09484	Q09484 caenorhabdi
40	80	11.3	564	10	Q8L7Z9	Q817z9 spinacia ol
41	79	11.2	568	16	Q9CKM1	Q9ckm1 pasteurella
42	79	11.2	618	5	Q22003	Q22003 caenorhabdi
43	79	11.2	638	16	Q83F64	Q83f64 coxiella bu
44	78.5	11.1	261	16	Q83KX1	Q83kx1 shigella fl
45	78.5	11.1	458	16	Q9X1C8	Q9x1c8 thermotoga

#### ALIGNMENTS

```
RESULT 1
Q7T224

ID Q7T224;

AC Q7T224;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE
     RTN4-C.
GN
     RTN4.
OS
     Gallus gallus (Chicken).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC
     Gallus.
    NCBI TaxID=9031;
OX
RN
     [1]
RP
     SEQUENCE FROM N.A.
RX
    MEDLINE=22715887; PubMed=12832288;
RA
     Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
RT
     "A reticular rhapsody: phylogenic evolution and nomenclature of the
RT
     RTN/Nogo gene family.";
RT.
     FASEB J. 17:1238-1247(2003).
DR
     EMBL; AY164737; AAP47312.1; -.
SQ
     SEQUENCE
               199 AA; 22293 MW; 07CF4E4EF2723251 CRC64;
  Query Match
                         65.2%; Score 460; DB 13; Length 199;
  Best Local Similarity
                         88.5%; Pred. No. 6.3e-39;
           92; Conservative
                                5; Mismatches 7; Indels
                                                               0;
                                                                   Gaps
                                                                           0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qy
                    Db
           1 MDSQPSGWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
Qv
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVOKYSNSALG 141
             Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESDVAVSEDLIQKYSSVVLG 104
RESULT 2
080W95
ID
                PRELIMINARY;
    Q80W95
                                 PRT;
                                        578 AA.
AC
    Q80W95;
DT
    01-JUN-2003 (TrEMBLrel. 24, Created)
DT
    01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
\mathsf{D}\mathbf{T}
    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
    Nogo-A (Fragment).
DΕ
    NOGO-A.
GN
OS
    Mus musculus (Mouse).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RA
    Tozaki H., Hirata T.;
RT
    "The partial sequence of mouse nogo-A cDNA clone#4109.";
    Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RL
DR
    EMBL; AB073672; BAC75974.1; -.
DR
    GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
FT
    NON TER
SO
    SEQUENCE
               578 AA; 63696 MW; 832670C171E4AC61 CRC64;
 Query Match
                         63.4%; Score 447; DB 11; Length 578;
```

```
Best Local Similarity
                        96.9%; Pred. No. 4.4e-37;
          93; Conservative
                              0; Mismatches 3; Indels
                                                            0;
                                                                       0:
                                                               Gaps
Qу
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
             388 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 447
Db
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
         448 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 483
RESULT 3
Q8K290
ID
    08K290
               PRELIMINARY;
                                PRT:
                                      639 AA.
AC
    08K290;
DT
    01-OCT-2002 (TrEMBLrel. 22, Created)
    01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT
    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT
DΕ
    Hypothetical protein.
GN
    RTN4.
OS
    Mus musculus (Mouse).
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RA
    Strausberg R.;
    Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RL
DR
    EMBL; BC032192; AAH32192.1; -.
DR
    MGD; MGI:1915835; Rtn4.
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
    GO; GO:0007399; P:neurogenesis; IDA.
DR
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
    PROSITE; PS50845; RETICULON; 1.
DR
KW
    Hypothetical protein.
SQ
    SEQUENCE
              639 AA; 70312 MW; 309A19DA37603F11 CRC64;
 Query Match
                        63.4%; Score 447; DB 11; Length 639;
 Best Local Similarity
                       96.9%; Pred. No. 4.9e-37;
          93; Conservative
                              0; Mismatches
                                                                       0;
                                              3;
                                                 Indels
                                                           0;
                                                               Gaps
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qy
               Db
         449 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 508
         106 QAIAKSDEGHPFRAYLESEVAISEELVOKYSNSALG 141
Qν
             Db
         509 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 544
RESULT 4
O8IUA4
    Q8IUA4
ID
               PRELIMINARY;
                                PRT:
                                      986 AA.
AC
    Q8IUA4;
    01-MAR-2003 (TrEMBLrel. 23, Created)
DT
```

```
DТ
     01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT
     01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
    RNT4 (RTN4).
DE
GN
    RTN4.
OS
    Homo sapiens (Human).
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
OC
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
    NCBI TaxID=9606;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RA
    Oertle T., van der Putten H., Schwab M.E.;
RT
    "Genomic Structure and Functional Characterization of the Promoter
RT
    Structures of Human and Mouse Nogo/Rtn-4.";
RL
    Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN
RP
    SEQUENCE FROM N.A.
    Oertle T., Schwab M.E.;
RΑ
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
    SEQUENCE FROM N.A.
RP
RA
    Van der Putten H.;
RL
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN
RP
    SEQUENCE FROM N.A.
RC
    TISSUE=Testis;
RX
    MEDLINE=22376540; PubMed=12488097;
    Oertle T., van der Putten H., Schwab M.E.;
RA
RT
    "Genomic Structure and Functional Characterization of the Promoter
RT
    Structures of Human and Mouse Nogo/Rtn-4.";
RL
    J. Mol. Biol. 325:299-323(2003).
DR
    EMBL; AY102285; AAM64244.1; -.
    EMBL; AY123245; AAM64249.1; -.
DR
    EMBL; AY123246; AAM64250.1; -.
DR
    EMBL; AY123247; AAM64251.1; -.
DR
DR
    EMBL; AY123248; AAM64252.1; -.
DR
    EMBL; AY123249; AAM64253.1; -.
DR
    EMBL; AY123250; AAM64254.1; -.
    GO; GO:0005783; C:endoplasmic reticulum; IEA.
    InterPro; IPR003388; Reticulon.
DR
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
SO
    SEQUENCE
               986 AA; 108449 MW; 0CDE8F647036415A CRC64;
 Query Match
                         63.4%; Score 447; DB 4; Length 986;
 Best Local Similarity
                         96.9%; Pred. No. 7.9e-37;
                              0; Mismatches 3; Indels
 Matches
           93; Conservative
                                                                          0;
                                                               0; Gaps
Qy
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
               Db
         796 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 855
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
         856 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 891
```

```
O8BGK7
ΙD
     Q8BGK7
                PRELIMINARY;
                                  PRT; 1046 AA.
AC
     08BGK7;
DТ
     01-MAR-2003 (TrEMBLrel. 23, Created)
     01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
     RTN4.
DE
GN
     RTN4.
OS
    Mus musculus (Mouse).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
     [1]
RΡ
     SEQUENCE FROM N.A.
RC
     STRAIN=129/SvcJ7, and 129SvcJ7;
RA
     Oertle T., van der Putten H., Schwab M.E.;
RT
     "Genomic Structure and Functional Characterization of the Promoter
RT
     Structures of Human and Mouse Nogo/Rtn-4.";
RL
     Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN
     [2]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7, and 129SvcJ7;
RA
    Oertle T., Schwab M.E.;
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
     [3]
RΡ
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7;
RA
    Van der Putten H.;
RL
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN
    SEQUENCE FROM N.A.
RP
    STRAIN=129SvcJ7;
RC
RA
    Van der Putten H., Mir A.;
ŔĹ
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR
    EMBL; AY102280; AAM73502.1; -.
DR
    EMBL; AY102286; AAM73507.1; -.
DR
    MGD; MGI:1915835; Rtn4.
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
    GO; GO:0007399; P:neurogenesis; IDA.
DR
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
    PROSITE; PS50845; RETICULON; 1.
DR
SQ
    SEQUENCE
               1046 AA; 114221 MW; 8CE2E2238ED51222 CRC64;
 Query Match
                         63.4%; Score 447; DB 11; Length 1046;
 Best Local Similarity
                         96.9%; Pred. No. 8.5e-37;
 Matches
          93; Conservative
                              0; Mismatches
                                                3; Indels
                                                              0; Gaps
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
             Db
         856 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 915
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
         916 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 951
```

```
RESULT 6
Q8BGM9
ID
    Q8BGM9
                PRELIMINARY;
                                PRT; 1162 AA.
AC
    Q8BGM9;
DT
    01-MAR-2003 (TrEMBLrel. 23, Created)
DT
    01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT
    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE
    RTN4.
GN
    RTN4.
OS
    Mus musculus (Mouse).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7, and 129SvcJ7;
    Oertle T., van der Putten H., Schwab M.E.;
RA
RT
    "Genomic Structure and Functional Characterization of the Promoter
    Structures of Human and Mouse Nogo/Rtn-4.";
RT
    Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
    [2]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7, and 129SvcJ7;
    Oertle T., Schwab M.E.;
RA
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
    [3]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7;
RA
    Van der Putten H.;
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
RP
    SEQUENCE FROM N.A.
    STRAIN=129SvcJ7;
RC
RA
    Van der Putten H., Mir A.;
RL
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR
    EMBL; AY102284; AAM73506.1; -.
    EMBL; AY102286; AAM73511.1; -.
DR
DR
    MGD; MGI:1915835; Rtn4.
DR
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
    GO; GO:0007399; P:neurogenesis; IDA.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
DR
             1162 AA; 126613 MW; 855697FBEE11781F CRC64;
SQ
    SEQUENCE
 Query Match
                         63.4%; Score 447; DB 11; Length 1162;
                        96.9%; Pred. No. 9.5e-37;
 Best Local Similarity
         93; Conservative
                              0: Mismatches
                                                3; Indels
                                                              0; Gaps
                                                                          0;
Qy
          46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
             Db
         972 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1031
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
             Db
        1032 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1067
```

```
RESULT 7
Q8BH78
ID
    Q8BH78
                PRELIMINARY;
                                 PRT;
                                        356 AA.
AC
    Q8BH78;
DT
    01-MAR-2003 (TrEMBLrel. 23, Created)
DT
    01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT
DE
GN
    RTN4.
OS
    Mus musculus (Mouse).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7, and 129SvcJ7;
RA
    Oertle T., van der Putten H., Schwab M.E.;
    "Genomic Structure and Functional Characterization of the Promoter
RT
RT
    Structures of Human and Mouse Nogo/Rtn-4.";
    Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7, and 129SvcJ7;
RA
    Oertle T., Schwab M.E.;
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
    [3]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=129/SvcJ7;
RA
    Van der Putten H.;
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
    [4]
RP
    SEQUENCE FROM N.A.
RC.
    STRAIN=129SvcJ7;
    Van der Putten H., Mir A.;
RA
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
DR
    EMBL; AY102281; AAM73503.1; -.
    EMBL; AY102286; AAM73508.1; -.
DR
    MGD; MGI:1915835; Rtn4.
DR
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
    GO; GO:0007399; P:neurogenesis; IDA.
DR
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
    SEQUENCE
             356 AA; 38403 MW; 4366C03BA9630B56 CRC64;
SQ
                         62.8%; Score 443; DB 11; Length 356;
 Query Match
 Best Local Similarity
                        98.9%; Pred. No. 6.5e-37;
                              0; Mismatches 1; Indels
 Matches
          92; Conservative
                                                              0; Gaps
                                                                          0;
Qу
          49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
             Db
         169 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 228
Qу
         109 AKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
              Db
         229 QKSDEGHPFRAYLESEVAISEELVQKYSNSALG 261
```

```
RESULT 8
Q8BHF5
TD
     Q8BHF5
                PRELIMINARY;
                                  PRT;
                                         375 AA.
AC
     Q8BHF5;
DT
     01-MAR-2003 (TrEMBLrel. 23, Created)
DT
     01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DΕ
     RTN4.
GN
     RTN4.
OS
     Mus musculus (Mouse).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
     NCBI TaxID=10090;
RN
     [1]
     SEQUENCE FROM N.A.
RP
     STRAIN=129/SvcJ7, and 129SvcJ7;
RC
RA
     Oertle T., van der Putten H., Schwab M.E.;
RT
     "Genomic Structure and Functional Characterization of the Promoter
RT
     Structures of Human and Mouse Nogo/Rtn-4.";
RL
     Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN
     [2]
RP
     SEQUENCE FROM N.A.
     STRAIN=129/SvcJ7, and 129SvcJ7;
RC
RA
     Oertle T., Schwab M.E.;
     Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RT.
RN
     [3]
RP
     SEQUENCE FROM N.A.
RC
     STRAIN=129/SvcJ7;
    Van der Putten H.:
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RT.
RN
RP
    SEQUENCE FROM N.A.
RC
     STRAIN=129SvcJ7;
    Van der Putten H., Mir A.;
RA
RL
     Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
     EMBL; AY102282; AAM73504.1; -.
DR
    EMBL; AY102286; AAM73509.1; -.
DR
DR
    MGD; MGI:1915835; Rtn4.
DR
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
    GO; GO:0007399; P:neurogenesis; IDA.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
     PROSITE; PS50845; RETICULON; 1.
DR
SQ
               375 AA; 40300 MW; 23D9EB19BE671AE6 CRC64;
     SEQUENCE
                         62.8%; Score 443; DB 11; Length 375;
 Query Match
 Best Local Similarity
                         98.9%; Pred. No. 6.9e-37;
 Matches
           92; Conservative
                                0; Mismatches
                                                 1; Indels
          49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
Qy
             Db
         188 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 247
         109 AKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
```

Qу

```
RESULT 9
096B16
ID
     Q96B16
                 PRELIMINARY;
                                   PRT;
                                          392 AA.
AC
     Q96B16;
DТ
     01-DEC-2001 (TrEMBLrel. 19, Created)
DТ
     01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT
     01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DΕ
     Hypothetical protein (RTN4).
GN
     RTN4.
OS
     Homo sapiens (Human).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
     NCBI TaxID=9606;
RN
RP
     SEQUENCE FROM N.A.
RC
     TISSUE=Kidney;
RA
     Strausberg R.;
RL
     Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE FROM N.A.
RA
     Oertle T., van der Putten H., Schwab M.E.;
RT
     "Genomic Structure and Functional Characterization of the Promoter
RT
     Structures of Human and Mouse Nogo/Rtn-4.";
RL
     Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE FROM N.A.
RA
     Oertle T., Schwab M.E.;
RL
     Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE FROM N.A.
RA
     Van der Putten H.;
     Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RL
RN
     [5]
RP
     SEQUENCE FROM N.A.
RX
    MEDLINE=22376540; PubMed=12488097;
     Oertle T., van der Putten H., Schwab M.E.;
RA
RT
     "Genomic Structure and Functional Characterization of the Promoter
RT
     Structures of Human and Mouse Nogo/Rtn-4.";
     J. Mol. Biol. 325:299-323(2003).
RL
DR
    EMBL; BC016165; AAH16165.1; -.
DR
    EMBL; AY102285; AAM64242.1; -.
DR
    EMBL; AY102278; AAM64247.1; -.
    GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR
    InterPro; IPR003388; Reticulon.
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
DR
KW
    Hypothetical protein.
    SEQUENCE 392 AA; 42274 MW; D7B2AA5E839E58AD CRC64;
SO
  Query Match
                          62.8%; Score 443; DB 4; Length 392;
 Best Local Similarity 98.9%; Pred. No. 7.3e-37;
          92; Conservative
 Matches
                                0; Mismatches
                                                 1; Indels
                                                                 0; Gaps
```

0;

```
205 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 264
Db
         109 AKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qy
              265 QKSDEGHPFRAYLESEVAISEELVQKYSNSALG 297
Db
RESULT 10
Q8K3G8
ID
     Q8K3G8
                PRELIMINARY;
                                PRT: 1163 AA.
AC
     Q8K3G8;
     01-OCT-2002 (TrEMBLrel. 22, Created)
DТ
DT
     01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DΕ
    Nogo-A.
GN
    RTN4.
    Mus musculus (Mouse).
OS
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC
OX
    NCBI TaxID=10090;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=BALB/c;
RA
    Jin W., Long M., Li R., Ju G.;
RΤ
    "Cloning and expression of the mouse Nogo-A protein.";
RL
    Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR
    EMBL; AY114152; AAM77068.1; -.
DR
    MGD; MGI:1915835; Rtn4.
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
DR
    GO; GO:0007399; P:neurogenesis; IDA.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
SO
    SEQUENCE
            1163 AA; 126691 MW;
                                  6B5F362799417EA4 CRC64;
 Query Match
                       61.9%; Score 436.5; DB 11; Length 1163;
 Best Local Similarity
                       95.9%; Pred. No. 1.1e-35;
 Matches
         93; Conservative
                             0; Mismatches
                                              3; Indels
Qу
         46 KDKVVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 104
              Db
         972 KTSVVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGV 1031
Qу
         105 IQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
            Db
        1032 IQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1068
RESULT 11
Q8K3G7
ID
    Q8K3G7
               PRELIMINARY;
                               PRT;
                                      357 AA.
AC
    08K3G7;
    01-OCT-2002 (TrEMBLrel. 22, Created)
    01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT
DE
    Nogo-B.
```

```
GN
     RTN4.
OS
     Mus musculus (Mouse).
OC.
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC
OX
     NCBI TaxID=10090;
RN
     [1]
     SEQUENCE FROM N.A.
RP
RC
     STRAIN=BALB/c;
     Jin W., Li R., Long M., Shen J., Ju G.;
RA
     "Cloning and expression of the mouse Nogo-B protein.";
RT
RT.
     Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR
     EMBL; AY114153; AAM77069.1; -.
     MGD; MGI:1915835; Rtn4.
DR
DR
     GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
     GO; GO:0007399; P:neurogenesis; IDA.
DR
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON; 1.
SO
     SEQUENCE
               357 AA; 38566 MW; 73BB3D17DFDBDF15 CRC64;
  Query Match
                          61.3%; Score 432.5; DB 11; Length 357;
  Best Local Similarity
                         97.9%; Pred. No. 7.7e-36;
  Matches
          92; Conservative
                                0; Mismatches
                                                  1; Indels
                                                                1; Gaps
                                                                            1;
Qу
           49 VVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQA 107
              169 VVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQA 228
Db
Qу
          108 IAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
              1 11111111111111111
Db
          229 IQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 262
RESULT 12
Q7YRW9
ID
     Q7YRW9
                PRELIMINARY;
                                  PRT;
                                         184 AA.
AC
     Q7YRW9;
DT
     01-OCT-2003 (TrEMBLrel. 25, Created)
     01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DF.
    RTN4w (Fragment).
GN
    RTN4.
OS
    Bos taurus (Bovine).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC
    Bovidae; Bovinae; Bos.
OX
    NCBI TaxID=9913;
RN
     [1]
RP
    SEQUENCE FROM N.A.
    MEDLINE=22715887; PubMed=12832288;
RX
RA
    Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
    "A reticular rhapsody: phylogenic evolution and nomenclature of the
RT
RΤ
    RTN/Nogo gene family.";
RL
    FASEB J. 17:1238-1247(2003).
    EMBL; AY164744; AAP47319.1; -.
DR
FT
    NON TER
                  1
                         1
               184 AA; 20671 MW; DE990E03BBAF84A1 CRC64;
SO
    SEQUENCE
```

```
Query Match
                         59.1%; Score 417; DB 6; Length 184;
  Best Local Similarity 96.6%; Pred. No. 1.4e-34;
  Matches
           86; Conservative
                              0; Mismatches
                                                3; Indels
                                                                         0;
                                                             0; Gaps
           53 LYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD 112
Qу
              Db
            1 LYWRDIKKTGVVFGXXLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSD 60
Qу
          113 EGHPFRAYLESEVAISEELVOKYSNSALG 141
              Db
           61 EGHPFRAYLESEVAISEELVQKYSNSALG 89
RESULT 13
Q9GM33
ID
     Q9GM33
                PRELIMINARY;
                                 PRT;
                                       179 AA.
AC
     Q9GM33;
     01-MAR-2001 (TrEMBLrel. 16, Created)
DТ
DТ
     01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT
DE
     Hypothetical protein.
     Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OS
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC.
OC
     Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC
     Cercopithecinae; Macaca.
OX
     NCBI TaxID=9541;
RN
     [1]
RP
     SEQUENCE FROM N.A.
RC
     TISSUE=Brain;
     Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA
RA
     Suzuki Y., Sugano S., Hashimoto K.;
     "Isolation of full-length cDNA clones from macaque brain cDNA
RT
RT
     libraries.";
     Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RL
    EMBL; AB049853; BAB16739.1; -.
DR
DR
    GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
    PROSITE; PS50845; RETICULON; 1.
DR
KW
    Hypothetical protein.
SO
    SEQUENCE 179 AA; 19949 MW; 5F8CD4383FEE9E02 CRC64;
  Query Match
                        54.8%; Score 386; DB 6; Length 179;
 Best Local Similarity
                        96.4%; Pred. No. 2e-31;
 Matches 81; Conservative
                              2; Mismatches
                                               1; Indels
                                                            0; Gaps
                                                                        0;
          58 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSDEGHPF 117
Qy
             Db
           1 MKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYRGVIQAIQKSDEGHPF 60
Qy
         118 RAYLESEVAISEELVQKYSNSALG 141
             111111111111111111111111
Db
          61 RAYLESEVAISEELVQKYSNSALG 84
```

```
O7T222
 ID
                 PRELIMINARY;
     O7T222
                                   PRT;
                                          214 AA.
AC
     Q7T222;
DT
      01-OCT-2003 (TrEMBLrel. 25, Created)
      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT
DΕ
     RTN4-M.
GN
     RTN4.
OS
     Carassius auratus (Goldfish).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC
OC
     Cyprinidae; Carassius.
OX
     NCBI TaxID=7957;
RN
     [1]
RP
     SEQUENCE FROM N.A.
RX
     MEDLINE=22715887; PubMed=12832288;
RA
     Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
     "A reticular rhapsody: phylogenic evolution and nomenclature of the
RT
RT
     RTN/Nogo gene family.";
     FASEB J. 17:1238-1247(2003).
RL
     EMBL; AY164754; AAP47329.1; -.
DR
     SEQUENCE 214 AA; 24063 MW; E9B603B043159DD9 CRC64;
SQ
  Query Match
                          54.5%; Score 384; DB 13; Length 214;
  Best Local Similarity
                         71.4%; Pred. No. 3.9e-31;
  Matches
           70; Conservative 20; Mismatches
                                                                0; Gaps
                                                  8; Indels
Qу
           43 KHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYK 102
              Db
           21 KHWKEQVVDLLYWRDLQRTGVVFGASLLLLLSLSVCSIISVISYVALALLSVTISFRIYK 80
Qy
          103 GVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
              1::|1: ||:::|||:::||
Db
           81 GILQAVQKSEDGHPFKMYLDKDIGISSELVQKYSDTAL 118
RESULT 15
Q9BQ59
ID
                PRELIMINARY;
     Q9BQ59
                                  PRT:
                                         199 AA.
     09B059;
     01-JUN-2001 (TrEMBLrel. 17, Created)
DT
     01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT
     01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DΕ
     Similar to reticulon 1.
OS
    Homo sapiens (Human).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
    NCBI TaxID=9606;
RN
    [1]
    SEQUENCE FROM N.A.
RP
RC
    TISSUE=Lung;
RA
    Strausberg R.;
RL
    Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR
    EMBL; BC003003; AAH03003.1; -.
DR
    EMBL; BC000314; AAH00314.1; -.
    GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR
ĎR
    GO; GO:0000786; C:nucleosome; IEA.
```

```
DR
     GO; GO:0005634; C:nucleus; IEA.
DR
     GO; GO:0003677; F:DNA binding; IEA.
    GO; GO:0007001; P:chromosome organization and biogenesis (sen. . .; IEA.
DR
     GO; GO:0006334; P:nucleosome assembly; IEA.
DR
DR
    InterPro; IPR001951; Histone H4.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS00047; HISTONE H4; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
    SEQUENCE 199 AA; 22642 MW; 7CFA44CC568DF6D8 CRC64;
SQ
 Query Match
                        49.4%; Score 348; DB 4; Length 199;
 Best Local Similarity 64.0%; Pred. No. 1.7e-27;
          64; Conservative 18; Mismatches 18; Indels
                                                            0; Gaps
                                                                        0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qу
                  Db
           1 MDCVWSNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATIS 60
Qу
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
             11111 1:11: 1:111111:1111 1: :1:1 :111::
Db
          61 FRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 100
```

Search completed: September 3, 2004, 16:08:56 Job time: 11.8622 secs

## GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2004, 15:58:54; Search time 1.8747 Seconds

(without alignments)

3916.307 Million cell updates/sec

Title: US-09-830-972-32

Perfect score: 705

Sequence: 1 QASGEAGVSCLRENFAVYSV.....ESEVAISEELVQKYSNSALG 141

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt 42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result		% Query				
No.	Score	Match	Length	DB	ID	Description
1	495	70.2	199	1	RTN4_MOUSE	Q99p72 mus musculu
2	447	63.4	1163	1	RTN4_RAT	Q9jk11 rattus norv
3	447	63.4	1192	1	RTN4_HUMAN	Q9nqc3 homo sapien
4	337	47.8	776	1	RTN1 HUMAN	Q16799 homo sapien
5	337	47.8	777	1	RTN1 RAT	Q64548 rattus norv
6	309	43.8	236	1	RTN3 HUMAN	095197 homo sapien
7	308	43.7	237	1	RTN3 MOUSE	Q9es97 mus musculu
8	214	30.4	545	1	RTN2 HUMAN	075298 homo sapien
9	197	27.9	471	1	RTN2 MOUSE	070622 mus musculu
10	77.5	11.0	243	1	T2RD MOUSE	Q9jka2 mus musculu
11	77.5	11.0	261	1	PHSC ECOLI	P77409 escherichia
12	75.5	10.7	246	1	T2R8 MOUSE	Q9jka0 mus musculu
13	75.5	10.7	545	1	G6PI HELPY	025781 helicobacte
14	74.5	10.6	545	1	G6PI HELPJ	Q9zk49 helicobacte
15	74	10.5	268	1	YC73 HAEIN	P44150 haemophilus
16	72	10.2	614	1	S6AC RABIT	P48055 oryctolagus
17	72	10.2	3511	1	MY15_MOUSE	Q9qzz4 mus musculu

```
346 1 3BHS VACCC
18
         71
              10.1
                                                              P21097 v 3 beta-hy
19
                        346 1 3BHS VACCV
         71
               10.1
                                                              P26670 v 3 beta-hy
20
         70
             9.9
                        960 1 GBR1 MOUSE
                                                              Q9wv18 mus musculu
21
         70
                9.9
                        961 1
                                 GBR1 HUMAN
                                                              Q9ubs5 homo sapien
         70
                        991 1
                                                              Q9z0u4 rattus norv
22
                9.9
                                 GBR1 RAT
                                                      Q9z0u4 rattus norv
P21309 photobacter
Q97s89 streptococc
Q19084 caenorhabdi
Q9jka3 mus musculu
P45239 haemophilus
P33059 variola vir
Q10358 schizosacch
                        315 1
23
         69
                                 LXD1 PHOLE
                9.8
24
         69
                9.8
                        398 1
                                 PGK STRPN
                       468 1
25
         69
             9.8
                                 YDBM CAEEL
26
       68.5
               9.7
                     238 1
                                 T2RA MOUSE
27
       68.5
              9.7
                       311 1
                                 HTRB HAEIN
28
       68.5
               9.7
                       325 1
                                 VP35 VARV
                                 YDB5 SCHPO
                       184 1
29
         68
              9.6
                       1447 1
30
         68
                9.6
                                 DCC HUMAN
                                                              P43146 homo sapien
                                 VP35 VACCC
31
       67.5
               9.6
                       324 1
                                                              P20497 vaccinia vi
                                                  P29111

028244 arcm

P34183 caenorhabe

P43754 haemophilus

Q96r17 homo sapien

Q00669 drosophila

059831 schizosacch

P08158 emericella

Q58107 methanococc

Q9jyml neisseria

P43439 enterococc

034996 bacillus :
32
       67.5
               9.6
                        503 1
                                 LEU1 BUCUM
                                                              Q9evh0 buchnera ap
33
       67.5
                9.6
                        877 1
                                 SULH SCHPO
34
        67
                9.5
                        175 1
                                 OLE2 BRANA
                                 SYS ARCFU
35
        67
                9.5
                       453 1
36
        67
                9.5
                        525 1
                                 SYH CAEEL
37
        67
                9.5
                       756 1
                                 RIR1 HAEIN
38
      66.5
                9.4
                     3174 1
                                 CHAC HUMAN
                     253 1
39
       66
              9.4
                                 ADH DROAD
40
        66
              9.4
                       537 1
                                 YCUB SCHPO
41
        66 9.4 548 1 AMDS EMENI
42
       65.5
              9.3
                       182 1 Y696 METJA
                        398 1
43
      65.5
              9.3
                                 ACK1 NEIMB
                        664 1 NTPI_ENTHR
880 1 DPO1_BACSU
44
       65.5
              9.3
45
       65.5
                9.3
```

## ALIGNMENTS

```
RESULT 1
RTN4 MOUSE
     RTN4 MOUSE
                    STANDARD;
                                   PRT;
TD
                                           199 AA.
AC
     Q99P72; Q9CTE3;
DT
     28-FEB-2003 (Rel. 41, Created)
     28-FEB-2003 (Rel. 41, Last sequence update)
     28-FEB-2003 (Rel. 41, Last annotation update)
DT
DΕ
     Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein).
GN
     RTN4 OR NOGO.
OS
     Mus musculus (Mouse).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
     NCBI TaxID=10090;
RN
     [1]
RP
     SEQUENCE FROM N.A.
RC
     STRAIN=3T3-L1; TISSUE=Adipocyte;
RA
     Coulson A.C., Craggs P.D., Morris N.J.;
RT
     "Mouse vp20/RTN4C cDNA.";
RL
     Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN 
RP
     SEQUENCE OF 170-199 FROM N.A.
RC
     STRAIN=C57BL/6J; TISSUE=Embryo;
     MEDLINE=21085660; PubMed=11217851;
RX
RA
     Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
```

```
RA
     Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA
     Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA
     Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
     Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA
RA
     Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA
     Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
     Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA
RA
     Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA
     Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA
     Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA
     Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
     Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA
RA
     Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA
     Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA
     Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
     Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA
     Hayashizaki Y.;
RA
RT
     "Functional annotation of a full-length mouse cDNA collection.";
RL
     Nature 409:685-690(2001).
CC
     -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC
         block the regeneration of the nervous central system in adults (By
CC
         similarity).
CC
    -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
CC
         similarity).
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
CC
         membrane of the endoplasmic reticulum through 2 putative
         transmembrane domains (By similarity).
CC
CC
     -!- ALTERNATIVE PRODUCTS:
CC
         Event=Alternative splicing; Named isoforms=1;
CC
          Comment=A number of isoforms may be produced;
CC
        Name=1;
CC
          IsoId=Q99P72-1; Sequence=Displayed;
CC
    -!- SIMILARITY: Contains 1 reticulon domain.
CC
     _______
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
CC
     use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
CC
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
     or send an email to license@isb-sib.ch).
CC
     -----
DR
    EMBL; AF326337; AAK08076.1; -.
DR
     EMBL; AK003859; -; NOT ANNOTATED CDS.
DR
    MGD; MGI:1915835; Rtn4.
    GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR
    GO; GO:0030176; C:integral to endoplasmic reticulum membrane; ISS.
DR
    GO; GO:0005635; C:nuclear membrane; ISS.
DR
DR
    GO; GO:0005515; F:protein binding; ISS.
    GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR
DR
    GO; GO:0030517; P:negative regulation of axon extension; ISS.
    GO; GO:0007399; P:neurogenesis; IDA.
DR
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON; 1.
KW
     Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT
    DOMAIN
                  1
                        25
                               CYTOPLASMIC (Potential).
```

```
FT
                 26
     TRANSMEM
                        50
                                 POTENTIAL.
FT
     DOMAIN
                 51
                       137
                                 LUMENAL (Potential).
     TRANSMEM
FT
                138
                       162
                                 POTENTIAL.
                       199
FT
     DOMAIN
                163
                                 CYTOPLASMIC (Potential).
FT
     DOMAIN
                 12
                       199
                                 RETICULON.
SO
     SEQUENCE
               199 AA; 22466 MW; 07BE5D580059ED9C CRC64;
                         70.2%; Score 495; DB 1; Length 199;
  Query Match
  Best Local Similarity
                         97.1%; Pred. No. 2.6e-41;
  Matches 101; Conservative
                               0; Mismatches
                                                 3; Indels
                                                               0; Gaps
                                                                           0;
          38 MDGQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 97
Qv
              Db
           1 MDDQKKRWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTIS 60
          98 FRIYKGVIQAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
Qу
              Db
          61 FRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 104
RESULT 2
RTN4 RAT
                   STANDARD;
     RTN4 RAT
                                  PRT; 1163 AA.
AC
     Q9JK11; Q9JK10; Q9R0D9; Q9WUE9; Q9WUF0;
DT
     28-FEB-2003 (Rel. 41, Created)
     28-FEB-2003 (Rel. 41, Last sequence update)
DT
     10-OCT-2003 (Rel. 42, Last annotation update)
DT
DE
    Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DΕ
    (Glut4 vesicle 20 kDa protein).
GN
    RTN4 OR NOGO.
OS
    Rattus norvegicus (Rat).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX
    NCBI TaxID=10116;
RN
    [1]
    SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.
RP
RC
    STRAIN=Sprague-Dawley; TISSUE=Adipocyte;
RX
    MEDLINE=99249816; PubMed=10231557;
    Morris N.J., Ross S.A., Neveu J.M., Lane W.S., Lienhard G.E.;
RA
     "Cloning and characterization of a 22 kDa protein from rat adipocytes:
RT
RT
     a new member of the reticulon family.";
RL
    Biochim. Biophys. Acta 1450:68-76(1999).
RN
    SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RΡ
RX
    MEDLINE=20129258; PubMed=10667796;
RA
    Chen M.S., Huber A.B., Van der Haar M.E., Frank M., Schnell L.,
    Spillmann A.A., Christ F., Schwab M.E.;
RA
RT
     "Nogo-A is a myelin-associated neurite outgrowth inhibitor and an
RT
    antigen for monoclonal antibody IN-1.";
RL
    Nature 403:434-439(2000).
RN
    SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RP
RC
    STRAIN=Wistar Kyoto; TISSUE=Vascular smooth muscle;
RA
     Ito T., Schwartz S.M.;
RT
     "Cloning of a member of the reticulon gene family in rat: one of two
RT
    minor splice variants.";
RL
    Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
```

```
RN
     [4]
RP
     FUNCTION.
RX
     MEDLINE=22033691; PubMed=12037567;
RA
     GrandPre T., Li S., Strittmatter S.M.;
RT
     "Nogo-66 receptor antagonist peptide promotes axonal regeneration.";
RL
    Nature 417:547-551(2002).
CC
     -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC
        block the regeneration of the nervous central system in adults (By
CC
        similarity).
CC
     -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
CC
        similarity).
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
CC
        membrane of the endoplasmic reticulum through 2 putative
        transmembrane domains (By similarity).
CC
CC
     -!- ALTERNATIVE PRODUCTS:
CC
        Event=Alternative splicing; Named isoforms=4;
CC
        Name=1; Synonyms=Nogo-A, NI-220-250;
CC
          IsoId=Q9JK11-1; Sequence=Displayed;
CC
        Name=2; Synonyms=Nogo-B, Foocen-M1;
CC
          IsoId=Q9JK11-2; Sequence=VSP 005658;
CC
        Name=3; Synonyms=Nogo-C, VP20;
CC
          IsoId=Q9JK11-3; Sequence=VSP 005656, VSP 005657;
CC
        Name=4; Synonyms=Foocen-M2;
CC
          IsoId=Q9JK11-4; Sequence=VSP 005659;
CC
    -!- TISSUE SPECIFICITY: Isoforms 1, 2 and 3 are present in optic
CC
        nerve, spinal cord and cerebral cortex. Isoforms 1 and 2 are
CC
        present in dorsal root ganglion, sciatic nerve and PC12 cells
CC
        after longer exposure. Isoforms 2 and 3 are detected in kidney,
        cartilage, skin, lung and spleen. Isoform 3 is expressed at high
CC
        level in skeletal muscle. In adult animals isoform 1 is expressed
CC
CC
        mainly in the nervous system.
CC
    -!- SIMILARITY: Contains 1 reticulon domain.
     ______
CC
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
CC
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
CC
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
CC
    ______
DR
    EMBL; AF051335; AAF01564.1; -.
DR
    EMBL; AJ242961; CAB71027.1; -.
    EMBL; AJ242962; CAB71028.1; -.
DR
DR
    EMBL; AJ242963; CAB71029.1; -.
    EMBL; AF132045; AAD31019.1; -.
DR
    EMBL; AF132046; AAD31020.1; -.
DR
    GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.
DR
DR
    GO; GO:0005635; C:nuclear membrane; ISS.
DR
    GO; GO:0005515; F:protein binding; ISS.
    GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR
    GO; GO:0030517; P:negative regulation of axon extension; ISS.
    InterPro; IPR003388; Reticulon.
DR
DR
    Pfam; PF02453; Reticulon; 1.
    PROSITE; PS50845; RETICULON; 1.
DR
    Endoplasmic reticulum; Alternative splicing; Transmembrane.
KW
FT
    DOMAIN
                  1
                       989
                                CYTOPLASMIC (Potential).
```

```
FT
     TRANSMEM
                 990
                       1010
                                  POTENTIAL.
FT
     DOMAIN
                       1104
                1011
                                  LUMENAL (Potential).
     TRANSMEM
FT
                1105
                       1125
                                  POTENTIAL.
FT
     DOMAIN
                1126
                       1163
                                  CYTOPLASMIC (Potential).
FT
     DOMAIN
                 976
                       1163
                                  RETICULON.
FT
     DOMAIN
                  33
                         46
                                  POLY-GLU.
FT
     DOMAIN
                  73
                         76
                                  POLY-ALA.
     DOMAIN
FT
                 140
                        145
                                  POLY-PRO.
FT
     VARSPLIC
                  1
                        964
                                  Missing (in isoform 3).
FΤ
                                  /FTId=VSP 005656.
FΤ
     VARSPLIC
                 965
                        975
                                  AVLSAELSKTS -> MDGQKKHWKDK (in isoform
\Gamma T
                                  3).
FΤ
                                  /FTId=VSP 005657.
FT
     VARSPLIC
                 173
                        975
                                  Missing (in isoform 2).
FT
                                  /FTId=VSP 005658.
FT
     VARSPLIC
                 192
                        975
                                  Missing (in isoform 4).
FT
                                  /FTId=VSP 005659.
FT
     CONFLICT
                1130
                       1131
                                  MISSING (IN REF. 3; AAD31020).
SO
     SEQUENCE
                1163 AA; 126386 MW; 8CB894B09E94F0B6 CRC64;
  Query Match
                          63.4%; Score 447; DB 1; Length 1163;
  Best Local Similarity
                          96.9%; Pred. No. 8.2e-36;
           93; Conservative
  Matches
                                 0; Mismatches
                                                  3; Indels
                                                                0; Gaps
                                                                            0:
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qу
                 Db
          973 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1032
Qу
          106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
              Db
         1033 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1068
RESULT 3
RTN4 HUMAN
     RTN4 HUMAN
                   STANDARD;
                                  PRT; 1192 AA.
     Q9NQC3; 094962; Q9BXG5; Q9H212; Q9H3I3; Q9UQ42; Q9Y293; Q9Y2Y7;
AC
AC
     Q9Y5U6;
DT
     28-FEB-2003 (Rel. 41, Created)
     28-FEB-2003 (Rel. 41, Last sequence update)
DT
     15-MAR-2004 (Rel. 43, Last annotation update)
DΕ
     Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DΕ
     (Neuroendocrine-specific protein) (NSP) (Neuroendocrine specific
DΕ
     protein C homolog) (RTN-x) (Reticulon 5) (My043 protein).
GN
     RTN4 OR NOGO OR ASY OR KIAA0886.
OS
     Homo sapiens (Human).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
OX
    NCBI TaxID=9606;
RN
     [1]
RP
     SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX
    MEDLINE=20129242; PubMed=10667780;
    Prinjha R., Moore S.E., Vinson M., Blake S., Morrow R., Christie G.,
RA
    Michalovich D., Simmons D.L., Walsh F.S.;
RA
     "Inhibitor of neurite outgrowth in humans.";
RT
    Nature 403:383-384(2000).
RL
RN
     [2]
```

```
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP
 RC
      TISSUE=Brain;
RX
     MEDLINE=21010696; PubMed=11126360;
     Tagami S., Eguchi Y., Kinoshita M., Takeda M., Tsujimoto Y.;
RA
RT
      "A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on
RT
      endoplasmic reticulum and reduces their anti-apoptotic activity.";
RL
     Oncogene 19:5736-5746(2000).
RN
RP
     SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX
     MEDLINE=20237542; PubMed=10773680;
     Yang J., Yu L., Bi A.D., Zhao S.-Y.;
     "Assignment of the human reticulon 4 gene (RTN4) to chromosome
RT
RT
     2p14-->2p13 by radiation hybrid mapping.";
RL
     Cytogenet. Cell Genet. 88:101-102(2000).
RN
     [4]
RP
     SEQUENCE FROM N.A. (ISOFORM 4).
RA
     Jin W.-L., Ju G.;
     "Developmentally-regulated alternative splicing in a novel Nogo-A.";
RT
RL
     Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN
     SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RP
RC
     TISSUE=Placenta, and Skeletal muscle;
RA
     Ito T., Schwartz S.M.;
RT
     "Cloning of a member of the reticulon gene family in human.";
     Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RL
RN
RP
     SEQUENCE FROM N.A. (ISOFORM 2).
RC
     TISSUE=Fibroblast;
RA
     Yutsudo M.;
RT
     "Isolation of a cell death-inducing gene.";
RL
     Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE FROM N.A. (ISOFORM 3).
RC
     TISSUE=Pituitary;
RA
     Song H., Peng Y., Zhou J., Huang Q., Dai M., Mao Y.M., Yu Y., Xu X.,
RA
     Luo B., Hu R., Chen J.;
RT
     "Human neuroendocrine-specific protein C (NSP) homolog gene.";
RL
     Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN
     SEQUENCE FROM N.A. (ISOFORM 3).
RP
RA
     Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA
     Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA
     Yu J., Han L.H.;
RT
     "Novel human cDNA clone with function of inhibiting cancer cell
RT
     growth.";
RL
     Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE FROM N.A. (ISOFORM 1).
RC
     TISSUE=Brain;
RX
     MEDLINE=99156230; PubMed=10048485;
RA
     Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
     Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RA
     "Prediction of the coding sequences of unidentified human genes. XII.
RT
     The complete sequences of 100 new cDNA clones from brain which code
RT
RT
     for large proteins in vitro.";
     DNA Res. 5:355-364(1998).
RL
RN
     [10]
```

```
SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RP
RC
     TISSUE=Brain, Ovary, Pancreas, Placenta, and Skeletal muscle;
RX
     MEDLINE=22388257; PubMed=12477932;
RA
     Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
     Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA
RA
     Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA
     Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA
     Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA
     Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA
     Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA
     Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
     Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA
     Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA
     Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA
RA
     Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
     Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA
RA
     Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
     Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA
     Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA
RA
     Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT
     "Generation and initial analysis of more than 15,000 full-length
RT
     human and mouse cDNA sequences.";
RL
     Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP
     SEQUENCE FROM N.A. (ISOFORM 3).
     MEDLINE=20499367; PubMed=11042152;
RX
RA
     Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA
     Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
     Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RA
RT
     "Cloning and functional analysis of cDNAs with open reading frames for
     300 previously undefined genes expressed in CD34+ hematopoietic
RT
RT
     stem/progenitor cells.";
RL
     Genome Res. 10:1546-1560(2000).
RN
     SEQUENCE OF 482-1192 FROM N.A. (ISOFORM 1/4).
RP
RC
     TISSUE=Brain;
RA
     Mao Y.M., Xie Y., Zheng Z.H.;
RL
     Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN
RP
     SEQUENCE OF 186-1192 FROM N.A. (ISOFORM 1).
RC
     TISSUE=Testis;
RA
     Sha J.H., Zhou Z.M., Li J.M.;
RL
     Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN
     [14]
     TOPOLOGY.
RP
RC
     TISSUE=Brain;
RX
     MEDLINE=20129259; PubMed=10667797;
RA
     GrandPre T., Nakamura F., Vartanian T., Strittmatter S.M.;
RT
     "Identification of the Nogo inhibitor of axon regeneration as a
RT
     Reticulon protein.";
RL
     Nature 403:439-444(2000).
RN
     [15]
RP
     FUNCTION.
     TISSUE=Brain;
RC
RX
    MEDLINE=21069055; PubMed=11201742;
RA
     Fournier A.E., Grandpre T., Strittmatter S.M.;
RT
     "Identification of a receptor mediating Nogo-66 inhibition of axonal
```

```
RT
     regeneration.";
     Nature 409:341-346(2001).
RL
RN
     [16]
RΡ
     REVIEW.
RX
     MEDLINE=21888956; PubMed=11891768;
     Ng C.E.L., Tang B.L.;
RA
RT
     "Nogos and the Nogo-66 receptor: factors inhibiting CNS neuron
RT
     regeneration.";
RL
     J. Neurosci. Res. 67:559-565(2002).
     -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC
CC
         block the regeneration of the nervous central system in adults.
CC
         Isoform 2 reduces the anti-apoptotic activity of Bcl-xl and Bcl-2.
CC
         This is likely consecutive to their change in subcellular
CC
         location, from the mitochondria to the endoplasmic reticulum,
CC
         after binding and sequestration.
     -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2.
CC
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC
         reticulum. Anchored to the membrane of the endoplasmic reticulum
CC
         through 2 putative transmembrane domains.
CC
     -!- ALTERNATIVE PRODUCTS:
CC
         Event=Alternative splicing; Named isoforms=4;
CC
         Name=1; Synonyms=RTN 4A, Nogo-A, RTN-xL;
CC
           IsoId=Q9NQC3-1; Sequence=Displayed;
CC
         Name=2; Synonyms=RTN 4B, Nogo-B, RTN-xS, Foocen-M;
CC
           IsoId=Q9NQC3-2; Sequence=VSP 005655;
CC
         Name=3; Synonyms=RTN 4C, Nogo-C, Foocen-S;
CC
           IsoId=Q9NQC3-3; Sequence=VSP_005652, VSP 005653;
CC
         Name=4;
CC
           IsoId=Q9NQC3-4; Sequence=VSP 005654;
     -!- TISSUE SPECIFICITY: Isoform 1 \overline{i}s specifically expressed in brain
CC
CC
         and testis and weakly in heart and skeletal muscle. Isoform 2 is
CC
        widely expressed excepted for the liver. Isoform 3 is expressed in
CC
        brain, skeletal muscle and adipocytes. Isoform 4 is testis-
CC
        specific.
CC
    -!- SIMILARITY: Contains 1 reticulon domain.
CC
    -!- CAUTION: Ref.11 sequence differs from that shown due to
CC
        frameshifts in positions 1149 and 1156.
CC
    CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
CC
CC
    modified and this statement is not removed. Usage by and for commercial
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
CC
DR
    EMBL; AJ251383; CAB99248.1; -.
DR
    EMBL; AJ251384; CAB99249.1; -.
DR
    EMBL; AJ251385; CAB99250.1; -.
    EMBL; AB040462; BAB18927.1; -.
DR
DR
    EMBL; AB040463; BAB18928.1; -.
DR
    EMBL; AF148537; AAG12176.1; -.
DR
    EMBL; AF148538; AAG12177.1; -.
DR
    EMBL; AF087901; AAG12205.1; -.
DR
    EMBL; AF320999; AAG40878.1; -.
DR
    EMBL; AF132047; AAD31021.1; -.
DR
    EMBL; AF132048; AAD31022.1; -.
```

```
EMBL; AB015639; BAA83712.1; -.
DR
     EMBL; AF077050; AAD27783.1; -.
DR
DR
     EMBL; AF177332; AAG17976.1; -.
     EMBL; AB020693; BAA74909.1; -.
DR
     EMBL; BC001035; AAH01035.1; -.
DR
     EMBL; BC007109; AAH07109.1; -.
DR
DR
     EMBL; BC014366; AAH14366.1; -.
  Query Match
                         63.4%; Score 447; DB 1; Length 1192;
  Best Local Similarity
                         96.9%; Pred. No. 8.4e-36;
           93; Conservative 0; Mismatches
  Matches
                                                             0; Gaps
                                                3; Indels
                                                                           0;
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
Qy
              1002 KTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 1061
Db
Qy
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSNSALG 141
              Db
         1062 QAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 1097
RESULT 4
RTN1 HUMAN
     RTN1 HUMAN
                   STANDARD;
                                  PRT;
                                        776 AA.
AC
     Q16799; Q16800; Q16801;
DT
     16-OCT-2001 (Rel. 40, Created)
DT
     16-OCT-2001 (Rel. 40, Last sequence update)
DT
     10-OCT-2003 (Rel. 42, Last annotation update)
DΕ
     Reticulon 1 (Neuroendocrine-specific protein).
GN
    RTN1 OR NSP.
OS
    Homo sapiens (Human).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
    NCBI TaxID=9606;
RN
     [1]
     SEQUENCE FROM N.A. (ISOFORMS RTN1-A; RTN1-B AND RTN1-C).
RP
RC
    TISSUE=Lung carcinoma;
RX
    MEDLINE=93293865; PubMed=7685762;
RA
     Roebroek A.J.M., Van de Velde H.J.K., Van Bokhoven A., Broers J.L.V.,
RA
     Ramaekers F.C.S., Van de Ven W.J.M.;
RT
     "Cloning and expression of alternative transcripts of a novel
RT
     neuroendocrine-specific gene and identification of its 135-kDa
RT
     translational product.";
RL
    J. Biol. Chem. 268:13439-13447(1993).
RN
    [2]
RP
    ALTERNATIVE SPLICING.
RX
    MEDLINE=96429995; PubMed=8833145;
RA
    Roebroek A.J.M., Ayoubi T.A.Y., Van de Velde H.J.K.,
    Schoenmakers E.F.P.M., Pauli I.G.L., Van de Ven W.J.M.;
RA
    "Genomic organization of the human NSP gene, prototype of a novel gene
RT
RT
    family encoding reticulons.";
RL
    Genomics 32:191-199(1996).
RN
    [3]
    TISSUE SPECIFICITY.
RP
RX
    MEDLINE=98228245; PubMed=9560466;
    Hens J., Nuydens R., Geerts H., Senden N.H., Van de Ven W.J.M.,
RA
RA
    Roebroek A.J., van de Velde H.J.K., Ramaekers F.C., Broers J.L.;
```

```
RT
      "Neuronal differentiation is accompanied by NSP-C expression.";
RL
      Cell Tissue Res. 292:229-237(1998).
CC
     -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC
         membrane trafficking in neuroendocrine cells.
CC
     -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
     -!- ALTERNATIVE PRODUCTS:
CC
CC
         Event=Alternative splicing; Named isoforms=3;
CC
         Name=RTN1-A; Synonyms=NSP-A;
CC
           IsoId=Q16799-1; Sequence=Displayed;
CC
         Name=RTN1-B; Synonyms=NSP-B;
CC
           IsoId=Q16799-2; Sequence=VSP 005644;
CC
         Name=RTN1-C; Synonyms=NSP-C;
CC
           IsoId=Q16799-3; Sequence=VSP 005645, VSP 005646;
CC
     -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL AND NEUROENDOCRINE TISSUES
CC
         AND CELL CULTURES DERIVED THEREFROM. EXPRESSION OF ISOFORM RTN1-C
CC
         IS STRONGLY CORRELATED WITH NEURONAL DIFFERENTIATION.
     -!- PTM: Isoforms RTN1-A and RTN1-B are phosphorylated.
CC
CC
     -!- SIMILARITY: Contains 1 reticulon domain.
     ______
CC
     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
     the European Bioinformatics Institute. There are no restrictions on its
CC
     use by non-profit institutions as long as its content is in no way
CC
     modified and this statement is not removed. Usage by and for commercial
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
CC
     or send an email to license@isb-sib.ch).
CC
DR
     EMBL; L10333; AAA59950.1; -.
DR
     EMBL; L10334; AAA59951.1; -.
DR
     EMBL; L10335; AAA59952.1; -.
DR
     PIR; A46583; A46583.
DR
     PIR; I60904; I60904.
DR
     Genew; HGNC:10467; RTN1.
DR
     MIM; 600865; -.
DR
     GO; GO:0030176; C:integral to endoplasmic reticulum membrane; TAS.
DR
     GO; GO:0004871; F:signal transducer activity; NAS.
     GO; GO:0030182; P:neuron differentiation; TAS.
DR
     GO; GO:0007165; P:signal transduction; NAS.
DR
DR
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON; 1.
     Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW
KW
     Phosphorylation.
FT
     TRANSMEM
                603
                       623
                                 POTENTIAL.
FT
    TRANSMEM
                726
                       746
                                 POTENTIAL.
FT
    DOMAIN
                589
                       776
                                 RETICULON.
FT
    DOMAIN
                609
                       612
                                 POLY-LEU.
FT
    VARSPLIC
                 1
                       420
                                 Missing (in isoform RTN1-B).
FT
                                 /FTId=VSP 005644.
FT
    VARSPLIC
                 1
                       568
                                 Missing (in isoform RTN1-C).
FT
                                 /FTId=VSP 005645.
FT
    VARSPLIC
             569
                       588
                                 GPGPLGPGAPPPLLFLNKQK -> MQATADSTKMDCVWSNW
FT
                                 KSQ (in isoform RTN1-C).
FT
                                 /FTId=VSP 005646.
     SEQUENCE 776 AA; 83617 MW; CA5B6232353096FE CRC64;
SO
 Query Match
                        47.8%; Score 337; DB 1; Length 776;
```

```
Best Local Similarity 67.4%; Pred. No. 3e-25;
   Matches 62; Conservative 16; Mismatches 14; Indels
                                                                0; Gaps
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
QУ
               586 KQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 645
Db
Qу
          106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
              11: 1:111111:111 1: :1:1 :111::
Db
          646 QAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 677
RESULT 5
RTN1 RAT
ID
     RTN1 RAT
                  STANDARD; PRT; 777 AA.
AC
     Q64548; Q64547;
DT
     16-OCT-2001 (Rel. 40, Created)
     16-OCT-2001 (Rel. 40, Last sequence update)
DT
     10-OCT-2003 (Rel. 42, Last annotation update)
DT
     Reticulon 1 (Neuroendocrine-specific protein) (S-rex).
DE
GN
     RTN1 OR NSP.
     Rattus norvegicus (Rat).
OS
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC
OX
     NCBI TaxID=10116;
RN
RP
     SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).
RC
     STRAIN=Wistar; TISSUE=Brain cortex;
RX
     MEDLINE=96386034; PubMed=8793864;
RA
     Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,
     Georgiev G.P., Buchman V.L.;
RA
RT
     "Intracellular compartmentalization of two differentially spliced s-
     rex/NSP mRNAs in neurons.";
RT
     Mol. Cell. Neurosci. 7:289-303(1996).
RL
CC
     -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC
        membrane trafficking in neuroendocrine cells.
     -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane (By
CC
CC
        similarity).
CC
    -!- ALTERNATIVE PRODUCTS:
CC
         Event=Alternative splicing; Named isoforms=2;
CC
         Name=RTN1-B; Synonyms=S-RexB;
CC
           IsoId=Q64548-1; Sequence=Displayed;
CC
        Name=RTN1-S; Synonyms=S-RexS;
CC
           IsoId=Q64548-2; Sequence=VSP 005647, VSP 005648;
     -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND
CC
CC
         PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS
CC
        HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.
CC
        EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL
CC
        TYPES.
CC
    -!- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE
CC
        HINDBRAIN AND IN Ell IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC
CC
        DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN
CC
        THE HINDBRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB
CC
        DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE
CC
        HINDBRAIN.
CC
    -!- SIMILARITY: Contains 1 reticulon domain.
CC
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
     use by non-profit institutions as long as its content is in no way
CC
     modified and this statement is not removed. Usage by and for commercial
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
CC
     or send an email to license@isb-sib.ch).
CC
DR
     EMBL; U17604; AAC53046.1; -.
DR
     EMBL; U17603; AAC53045.1; -.
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
DR
     PROSITE; PS50845; RETICULON; 1.
     Endoplasmic reticulum; Alternative splicing; Transmembrane.
KW
FT
     TRANSMEM
                 604
                        624 POTENTIAL.
FT
     TRANSMEM
                 727
                        747
                                 POTENTIAL.
FT
     DOMAIN
                 590
                     777
                                RETICULON.
                 610
FT
     DOMAIN
                        613
                                 POLY-LEU.
FT
     VARSPLIC
                 1
                        569
                                 Missing (in isoform RTN1-S).
FT
                                 /FTId=VSP 005647.
FT
     VARSPLIC 570
                        589
                                 IPGPLGSDLVPPLPFFNKQK -> MQATADSTKMDCVWSNW
FT
                                 KSQ (in isoform RTN1-S).
FT
                                 /FTId=VSP 005648.
     SEQUENCE 777 AA; 83001 MW; AF7479C50F28D0AC CRC64;
SQ
  Query Match 47.8%; Score 337; DB 1; Length 777; Best Local Similarity 67.4%; Pred. No. 3e-25;
  Matches
          62; Conservative 16; Mismatches 14; Indels
                                                                            0;
Qу
           46 KDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVI 105
              Db
          587 KQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVL 646
Qу
         106 QAIAKSDEGHPFRAYLESEVAISEELVQKYSN 137
              11: 1:111111:111 | 1: :|:| :|1|::
Db
          647 QAVQKTDEGHPFKAYLELEITLSQEQIQKYTD 678
RESULT 6
RTN3 HUMAN
ID
     RTN3 HUMAN
                   STANDARD;
                              PRT;
                                         236 AA.
AC
    095197;
DT
     16-OCT-2001 (Rel. 40, Created)
DT
     16-OCT-2001 (Rel. 40, Last sequence update)
DT
    10-OCT-2003 (Rel. 42, Last annotation update)
DΕ
    Reticulon protein 3 (Neuroendocrine-specific protein-like 2) (NSP-like
DE
    protein II) (NSPLII).
GN
    RTN3 OR NSPL2.
OS
    Homo sapiens (Human).
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
OC
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
    NCBI TaxID=9606;
RN
RP
    SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC
    TISSUE=Retina;
RX
    MEDLINE=99265974; PubMed=10331947;
RA
    Moreira E.F., Jaworski C.J., Rodriguez I.R.;
```

```
RT
      "Cloning of a novel member of the reticulon gene family (RTN3): gene
      structure and chromosomal localization to 11q13.";
RT
RL
     Genomics 58:73-81(1999).
RN
     SEQUENCE FROM N.A.
RP
RA
     Huang X., Zhou Y., Du G., Yuan J., Qiang B.;
RT
     "Cloning and expression analysis of a cDNA encoding a novel
     neuroendocrine-specific protein-like protein 1: NSPL1.";
RT
     Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RL
RN
     [3]
     SEQUENCE FROM N.A.
RP
RC
     TISSUE=Brain, Eye, and Lymph;
RX
     MEDLINE=22388257; PubMed=12477932;
RA
     Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
     Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA
     Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA:
RA
     Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
     Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA
     Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA
     Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA
     Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA
     Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA
RA
     Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA
     Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA
     Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
     Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA
RA
     Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA
     Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA
     Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA
     Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
     "Generation and initial analysis of more than 15,000 full-length
RT
RT
     human and mouse cDNA sequences.";
RL
     Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC
         reticulum (Potential).
CC
     -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
CC
         BRAIN. THREE TIMES MORE ABUNDANT IN MACULA THAN IN PERIPHERAL
CC
         RETINA.
CC
     -!- SIMILARITY: Contains 1 reticulon domain.
CC
     ______
     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
CC
     or send an email to license@isb-sib.ch).
CC
     DR
    EMBL; AF059524; AAC99319.1; -.
    EMBL; AF059529; AAD20951.1; -.
DR
DR
    EMBL; AF059525; AAD20951.1; JOINED.
DR
    EMBL; AF059526; AAD20951.1; JOINED.
DR
    EMBL; AF059527; AAD20951.1; JOINED.
    EMBL; AF059528; AAD20951.1; JOINED.
DR
    EMBL; AF119297; AAD26810.1; -.
DR
    EMBL; BC000634; AAH00634.1; -.
DR
    EMBL; BC010556; AAH10556.1; -.
DR
```

```
DR
     EMBL; BC011394; AAH11394.1; -.
     EMBL; BC022993; AAH22993.1; -.
DR
DR
     Genew; HGNC:10469; RTN3.
DR
     MIM; 604249; -.
DR
     GO; GO:0005615; C:extracellular space; TAS.
DR
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON; 1.
KW
     Transmembrane; Endoplasmic reticulum.
FT
     TRANSMEM
                  68
                         88
                                  POTENTIAL.
FT
     TRANSMEM
                 177
                        197
                                  POTENTIAL.
FT
     DOMAIN
                  48
                        236
                                  RETICULON.
SQ
     SEQUENCE
                236 AA; 25609 MW; DDC6A4544ABCDFB7 CRC64;
  Query Match
                          43.8%; Score 309; DB 1; Length 236;
  Best Local Similarity
                          60.9%; Pred. No. 4.7e-23;
  Matches
           56; Conservative 18; Mismatches
                                                 18; Indels
                                                                 0; Gaps
                                                                             0;
Qу
           49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
              Db
           48 VHDLIFWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAV 107
Qу
          109 AKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
               11:1111:11: :: :1 1
                                       1 1:1:
Db
          108 QKSEEGHPFKAYLDVDITLSSEAFHNYMNAAM 139
RESULT 7
RTN3 MOUSE
ID
     RTN3 MOUSE
                    STANDARD;
                                   PRT;
                                          237 AA.
     Q9ES97;
AC
     16-OCT-2001 (Rel. 40, Created)
DT
DT
     16-OCT-2001 (Rel. 40, Last sequence update)
DT
     15-MAR-2004 (Rel. 43, Last annotation update)
DΕ
     Reticulon protein 3.
GN
     RTN3.
OS
     Mus musculus (Mouse).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
     NCBI TaxID=10090;
RN
     [1]
RΡ
     SEQUENCE FROM N.A.
RA
     Huang X., Zhou Y., Qiang H., Yuan J., Qiang B.;
RT
     "Cloning and expression profile of a novel mouse cDNA encoding a human
RT
     RTN3 homolog.";
RL
     Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN
     [2]
RP
     SEQUENCE FROM N.A.
RC
    TISSUE=Eye;
    MEDLINE=22388257; PubMed=12477932;
RX
RA
    Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
    Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA
RA
    Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
    Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA
    Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA
    Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA
     Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA
```

```
RA
     Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA
     Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
     Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA
     Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA
RA
     Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
     Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA
RA
     Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA
     Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
     Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA
     Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA
     "Generation and initial analysis of more than 15,000 full-length
RT
RT
     human and mouse cDNA sequences.";
RL
     Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC
         reticulum (Potential).
CC
     -!- SIMILARITY: Contains 1 reticulon domain.
CC
     CC
     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
CC
     use by non-profit institutions as long as its content is in no way
CC
     modified and this statement is not removed. Usage by and for commercial
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
     or send an email to license@isb-sib.ch).
CC
DR
     EMBL; AF195940; AAG31360.1; -.
DR
     EMBL; BC014697; AAH14697.1; -.
DR
     MGD; MGI:1339970; Rtn3.
DR
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON: 1.
KW
     Transmembrane; Endoplasmic reticulum.
FT
     TRANSMEM
                 69
                        89
                                POTENTIAL.
FT
     TRANSMEM
                167
                       187
                                POTENTIAL.
FT
     DOMAIN
                 49
                       237
                                RETICULON.
               237 AA; 25428 MW; EB60A0A7AC45F0DE CRC64;
SO
     SEQUENCE
                         43.7%; Score 308; DB 1; Length 237;
  Query Match
  Best Local Similarity
                         59.8%; Pred. No. 6e-23;
  Matches
          55; Conservative 19; Mismatches 18; Indels
                                                              0; Gaps
                                                                          0;
          49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
Qу
             Db
          49 VHDLIFWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRVYKSVIQAV 108
         109 AKSDEGHPFRAYLESEVAISEELVQKYSNSAL 140
Qу
              11:11/11:11: :: :1 1
                                     1 1:1:
Db
         109 QKSEEGHPFKAYLDVDITLSSEAFHNYMNAAM 140
RESULT 8
RTN2 HUMAN
    RTN2 HUMAN
                   STANDARD;
                                 PRT;
                                        545 AA.
AC
    075298; 060509;
DT
    16-OCT-2001 (Rel. 40, Created)
DT
    16-OCT-2001 (Rel. 40, Last sequence update)
DT
    10-OCT-2003 (Rel. 42, Last annotation update)
```

```
DE
    Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE
    protein 1) (NSPLI).
GN
    RTN2 OR NSPL1.
OS
    Homo sapiens (Human).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX
    NCBI TaxID=9606;
RN
RP
    SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC
    TISSUE=Lung carcinoma;
RX
    MEDLINE=98360096; PubMed=9693037;
RA
    Roebroek A.J.M., Contreras B., Pauli I.G.L., Van de Ven W.J.M.;
RT
    "cDNA cloning, genomic organization, and expression of the human RTN2
RT
    gene, a member of a gene family encoding reticulons.";
RL
    Genomics 51:98-106(1998).
RN
RP
    SEQUENCE OF 108-545 FROM N.A. (ISOFORM RTN2-B).
RC
    TISSUE=Brain;
RX
    MEDLINE=98191726; PubMed=9530622;
RA
    Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT
    "Molecular cloning of a novel mouse gene with predominant muscle and
RT
    neural expression.";
RL
    Mamm. Genome 9:274-282(1998).
    -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC
CC
        reticulum (Potential).
    -!- ALTERNATIVE PRODUCTS:
CC
CC
        Event=Alternative splicing; Named isoforms=2;
CC
        Name=RTN2-A;
CC
          IsoId=075298-1; Sequence=Displayed;
CC
          Note=Isoform RTN2-C is produced by alternative initiation at
CC
          Met-341 of isoform RTN2-A;
CC
        Name=RTN2-B;
CC
          IsoId=075298-2; Sequence=VSP 005649;
CC
        Event=Alternative initiation;
          Comment=2 isoforms, RTN2-A (shown here) and RTN2-C, are produced
CC
CC
          by alternative initiation at Met-1 and Met-341;
    -!- TISSUE SPECIFICITY: ISOFORM RTN2-C IS HIGHLY EXPRESSED IN SKELETAL
CC
CC
        MUSCLE.
CC
    -!- SIMILARITY: Contains 1 reticulon domain.
CC
    CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
    _______
CC
    EMBL; AF004222; AAC32542.1; -.
DR
DR
    EMBL; AF004223; AAC32543.1; -.
DR
    EMBL; AF004224; AAC32544.1; -.
DR
    EMBL; AF038540; AAC14910.1; -.
    Genew; HGNC:10468; RTN2.
DR
DR
    MIM; 603183; -.
DR
    GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.
DR
    GO; GO:0004871; F:signal transducer activity; NAS.
DR
    GO; GO:0007165; P:signal transduction; NAS.
```

```
DR
     InterPro; IPR003388; Reticulon.
DR
     Pfam; PF02453; Reticulon; 1.
DR
     PROSITE; PS50845; RETICULON; 1.
KW
     Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW
     Alternative initiation.
FT
                                  RETICULON PROTEIN 2, ISOFORM RTN2-A.
                        545
     CHAIN
                  1
                        545
                                  RETICULON PROTEIN 2, ISOFORM RTN2-C.
FT
     CHAIN
                 341
     INIT MET
                        341
FT
                 341
                                  FOR ISOFORM RTN2-C.
     TRANSMEM
                        388
FT
                 368
                                  POTENTIAL.
FT
     TRANSMEM
                 463
                        483
                                  POTENTIAL.
FT
                 345
                        545
                                  RETICULON.
     DOMAIN
FT
     VARSPLIC
                272
                        344
                                 Missing (in isoform RTN2-B).
FT
                                  /FTId=VSP 005649.
SQ
     SEQUENCE
               545 AA; 59263 MW; 971FD2F909E1E9E6 CRC64;
  Query Match
                          30.4%; Score 214; DB 1; Length 545;
                                 Pred. No. 2.2e-13;
  Best Local Similarity
                          46.7%;
            42; Conservative 21; Mismatches
                                                 27; Indels
                                                                0; Gaps
                                                                             0;
           48 KVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQA 107
Qy
              11 1111:1::::1111
                                 Db
          344 KVADLLYWKDTRTSGVVFTGLMVSLLCLLHFSIVSVAAHLALLLLCGTISLRVYRKVLQA 403
          108 IAKSDEGHPFRAYLESEVAISEELVQKYSN 137
Qy
              :: | :||:|||: :: :: | :: |:
Db
          404 VHRGDGANPFQAYLDVDLTLTREQTERLSH 433
RESULT 9
RTN2 MOUSE
     RTN2 MOUSE
                    STANDARD;
                                  PRT;
                                         471 AA.
AC
     070622; 070620;
DT
     16-OCT-2001 (Rel. 40, Created)
DT
     16-OCT-2001 (Rel. 40, Last sequence update)
DT
     15-MAR-2004 (Rel. 43, Last annotation update)
DE
     Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE
     protein 1) (NSPLI).
    RTN2 OR NSPL1.
GN
OS
    Mus musculus (Mouse).
OC
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
     NCBI TaxID=10090;
RN
RP
     SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
RC
     STRAIN=FVB/N, and 129/Sv; TISSUE=Cerebellum, and Skeletal muscle;
RX
    MEDLINE=98191726; PubMed=9530622;
     Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RA
RT
     "Molecular cloning of a novel mouse gene with predominant muscle and
RT
     neural expression.";
RL
    Mamm. Genome 9:274-282(1998).
RN
RP
     SEQUENCE FROM N.A. (ISOFORM 1).
RC
     TISSUE=Retina;
RX
    MEDLINE=22388257; PubMed=12477932;
RA
     Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA
     Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA
     Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA
     Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA
RA
     Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
     Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA
RA
     Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA
     Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA
     Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA
     Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA
     Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
     Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA
RA
     Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA
     Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA
     Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
     Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA
     "Generation and initial analysis of more than 15,000 full-length
RT
RT
     human and mouse cDNA sequences.";
RL
     Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC
     -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
CC
         (Potential).
CC
     -!- ALTERNATIVE PRODUCTS:
CC
         Event=Alternative splicing; Named isoforms=2;
CC
         Name=1; Synonyms=Brain;
CC
           IsoId=070622-1; Sequence=Displayed;
CC
         Name=2; Synonyms=Muscle;
CC
           IsoId=070622-2; Sequence=VSP 005650, VSP 005651;
CC
     -!- TISSUE SPECIFICITY: Expressed predominantly in neural and muscular
CC
         tissues.
CC
     -!- SIMILARITY: Contains 1 reticulon domain.
CC
     ______
     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
     the European Bioinformatics Institute. There are no restrictions on its
CC
     use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
CC
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
     or send an email to license@isb-sib.ch).
CC
DR
    EMBL; AF038537; AAC14906.1; -.
    EMBL; AF038537; AAC14907.1; -.
DR
    EMBL; AF038538; AAC14908.1; -.
DR
DR
    EMBL; AF038539; AAC14909.1; -.
DR
    EMBL; AF093624; AAD13195.1; -.
DR
    EMBL; BC031370; AAH31370.1; -.
DR
    MGD; MGI:107612; Rtn2.
DR
    InterPro; IPR003388; Reticulon.
DR
    Pfam; PF02453; Reticulon; 1.
DR
    PROSITE; PS50845; RETICULON; 1.
    Endoplasmic reticulum; Alternative splicing; Transmembrane.
KW
FT
    TRANSMEM
                295
                       315
                                 POTENTIAL.
FT
    DOMAIN
                272
                       471
                                 RETICULON.
FT
    VARSPLIC
                       267
                                 Missing (in isoform 2).
FT
                                 /FTId=VSP 005650.
FT
    VARSPLIC
                268
                       271
                                 PLLL -> MGSK (in isoform 2).
FT
                                 /FTId=VSP 005651.
SQ
    SEQUENCE
               471 AA; 51346 MW; 9BBD8F372CF63AD3 CRC64;
 Query Match
                         27.9%; Score 197; DB 1; Length 471;
```

```
Best Local Similarity
                          44.3%; Pred. No. 8.6e-12;
           39; Conservative 20; Mismatches
                                                29; Indels
                                                                0; Gaps
                                                                            0;
Qу
           49 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 108
              Db
          272 VADLLYWKDTRTSGAVFTGLMASLLCLLHFSIVSVAAHLALLGLCATISLRVYRKVLOAV 331
Qу
          109 AKSDEGHPFRAYLESEVAISEELVOKYS 136
               : | :||:|||: :: :: | :: |
Db
          332 HRGDGTNPFQAYLDMDLTLTREOTERLS 359
RESULT 10
T2RD MOUSE
ΙD
     T2RD MOUSE
                   STANDARD;
                                   PRT;
                                         243 AA.
AC
     09JKA2;
     10-OCT-2003 (Rel. 42, Created)
DT
DΤ
     10-OCT-2003 (Rel. 42, Last sequence update)
DT
     10-OCT-2003 (Rel. 42, Last annotation update)
     Taste receptor type 2 member 13 (T2R13) (Taste receptor family B
DE
DΕ
    member 3) (TRB3) (Fragment).
GN
    TAS2R13.
OS
    Mus musculus (Mouse).
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC
OC
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX
    NCBI TaxID=10090;
RN
     [1]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=DBA/2J;
RX
    MEDLINE=20227309; PubMed=10766242;
RA
    Matsunami H., Montmayeur J.-P., Buck L.B.;
     "A family of candidate taste receptors in human and mouse.";
RТ
RL
    Nature 404:601-604(2000).
RN
    [2]
RP
    REVIEW.
RX
    MEDLINE=22135574; PubMed=12139982;
RA
    Montmayeur J.-P., Matsunami H.;
    "Receptors for bitter and sweet taste.";
RL
    Curr. Opin. Neurobiol. 12:366-371(2002).
RN
    [3]
RP
    REVIEW.
RX
    MEDLINE=21634924; PubMed=11696554;
RA
    Margolskee R.F.;
    "Molecular mechanisms of bitter and sweet taste transduction.";
RT
    J. Biol. Chem. 277:1-4(2002).
RL
RN
    [4]
    REVIEW.
RΡ
RX
    MEDLINE=22469025; PubMed=12581520;
RA
    Zhang Y., Hoon M.A., Chandrashekar J., Mueller K.L., Cook B., Wu D.,
RA
    Zuker C.S., Ryba N.J.;
RT
    "Coding of sweet, bitter, and umami tastes: different receptor cells
RT
    sharing similar signaling pathways.";
RL
    Cell 112:293-301(2003).
CC
    -!- FUNCTION: Receptor that may play a role in the perception of
CC
        bitterness and is gustducin-linked. May play a role in sensing the
CC
        chemical composition of the gastrointestinal content. The activity
CC
        of this receptor may stimulate alpha gustducin, mediate PLC-beta-2
```

```
CC
        activation and lead to the gating of TRPM5.
CC
    -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
    -!- TISSUE SPECIFICITY: Expressed in subsets of taste receptor cells
CC
        of the tongue and palate epithelium and exclusively in gustducin-
CC
        positive cells.
CC
    -!- MISCELLANEOUS: Most taste cells may be activated by a limited
CC
        number of bitter compounds; individual taste cells can
CC
        discriminate among bitter stimuli.
CC
    -!- SIMILARITY: Belongs to family T2R of G-protein coupled receptors.
    ______
CC
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
CC
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
CC
    DR
    EMBL; AF247733; AAF64510.1; -.
DR
    MGD; MGI:1890148; Tas2r13.
DR
    InterPro; IPR007960; TAS2R.
    Pfam; PF05296; TAS2R; 1.
KW
    Receptor; G-protein coupled receptor; Transmembrane.
FT
    NON TER
                1
                       1
                       12
FT
    DOMAIN
                <1
                               CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
                13
                       33
                               2 (POTENTIAL).
FT
                34
                      54
    DOMAIN
                               EXTRACELLULAR (POTENTIAL).
                55
                      75
FT
    TRANSMEM
                               3 (POTENTIAL).
                76
                      99
FT
    DOMAIN
                               CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
               100
                    120
                               4 (POTENTIAL).
FT
    DOMAIN
               121 150
                               EXTRACELLULAR (POTENTIAL).
                    171
195
FT
    TRANSMEM
               151
                               5 (POTENTIAL).
FT
    DOMAIN
               172
                               CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
               196 216
                               6 (POTENTIAL).
FT
    DOMAIN
               217
                     222
                               EXTRACELLULAR (POTENTIAL).
FT
    TRANSMEM
               223 >243
                               7 (POTENTIAL).
FT
    CARBOHYD
               128
                     128
                               N-LINKED (GLCNAC. . .) (POTENTIAL).
FT
    NON TER
               243
                      243
    SEQUENCE
              243 AA; 28110 MW; D8AD14AF95B9E0B2 CRC64;
SQ
 Query Match
                        11.0%; Score 77.5; DB 1; Length 243;
 Best Local Similarity 27.1%; Pred. No. 2;
 Matches 32; Conservative 18; Mismatches 47; Indels
                                                           21; Gaps
QУ
          17 VYSVSVGMHNLLLLEGRSWQEMDGQKKH---WKDKVVDLLYWRDIKKTGVVFGASLFLLL 73
            :|| : :|: :|: | | : : |: |
                                                          | |:|| |
Db
          37 LYSALMTTRKVLIIFNNSWTVIN----HFNIWLATCLSIFYFLKIAN----FSNSIFLSL 88
Qу
          74 SLTVFSIVSVTAYIALALLSV-----TISFRIYKGVIQAIAKSDEG-HPFRAYL 121
               | ::|||| ::| || |
                                         11 : 11
                                                     : 11
                                                          1 1 : 1
          89 RWRVKTVVSVTLMMSLLLLFVNVLVINTFIVISVDVYKVNTSYSSHSDNNLHISRIFL 146
RESULT 11
PHSC ECOLI
ID
    PHSC ECOLI
                  STANDARD; PRT;
                                       261 AA.
AC
    P77409;
```

```
DT
     01-NOV-1997 (Rel. 35, Created)
DT
     01-NOV-1997 (Rel. 35, Last sequence update)
     16-OCT-2001 (Rel. 40, Last annotation update)
DT
DE
     PhsC protein homolog.
     YDHU OR B1670.
GN
     Escherichia coli.
OS
OC
     Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
     Enterobacteriaceae; Escherichia.
     NCBI_TaxID=562;
OX
RN
     [1]
     SEQUENCE FROM N.A.
RP
RC
     STRAIN=K12 / MG1655;
RX
     MEDLINE=97426617; PubMed=9278503;
     Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA
RA
     Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA
     Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA
     Mau B., Shao Y.;
RT
     "The complete genome sequence of Escherichia coli K-12.";
RL
     Science 277:1453-1474(1997).
RN
RP
     SEQUENCE FROM N.A.
RC
     STRAIN=K12 / MG1655;
    MEDLINE=97175536; PubMed=9023191;
RX
     Hensel M., Shea J.E., Baeumler A.J., Gleeson C., Blattner F.R.,
RA
RA
     Holden D.W.;
RT
     "Analysis of the boundaries of Salmonella pathogenicity island 2 and
RT
     the corresponding chromosomal region of Escherichia coli K-12.";
RL
     J. Bacteriol. 179:1105-1111(1997).
RN
RP
     SEQUENCE FROM N.A.
RC
     STRAIN=K12;
RX
    MEDLINE=97251357; PubMed=9097039;
RA
    Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA
     Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA
     Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
    Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA
RA
     Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,
     Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA
     Yamamoto Y., Horiuchi T.;
RA
RT
     "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT
     corresponding to the 28.0-40.1 min region on the linkage map.";
RL
     DNA Res. 3:363-377(1996).
CC
    -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC
         (Potential).
CC
    -!- SIMILARITY: TO S.TYPHIMURIUM PHSC.
CC
     ______
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
CC
     use by non-profit institutions as long as its content is in no way
     modified and this statement is not removed. Usage by and for commercial
CC
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
     or send an email to license@isb-sib.ch).
CC
     _____
DR
    EMBL; AE000262; AAC74740.1; -.
DR
     EMBL; U68703; AAB47946.1; -.
DR
     EMBL; D90810; BAA15442.1; -.
```

```
DR
     PIR; F64924; F64924.
DR
     EcoGene; EG13955; ydhU.
DR
     InterPro; IPR000516; Ni hydr CytB.
     Pfam; PF01292; Ni hydr CYTB; 1.
DR
     Transmembrane; Inner membrane; Complete proteome.
KW
FT
     TRANSMEM
               25
                       45
                                POTENTIAL.
     TRANSMEM
                 81
                       101
FT
                                 POTENTIAL.
     TRANSMEM
               108
                      128
FT
                                 POTENTIAL.
FT
     TRANSMEM
               182
                       202
                                 POTENTIAL.
FT
     TRANSMEM
                224
                       244
                                 POTENTIAL.
SO
     SEQUENCE 261 AA; 29583 MW; 65CF1A45691A0AF3 CRC64;
  Query Match
                         11.0%; Score 77.5; DB 1; Length 261;
  Best Local Similarity
                         27.6%; Pred. No. 2.2;
  Matches
          29; Conservative 16; Mismatches
                                               43; Indels
                                                              17; Gaps
          22 VGMHNLLLLEGRSWQEMD-GQKKHWKDKVVDLLYWRDIKKTGVVFGASLFLLL----SLT 76
Qy
             | ||:||
          44 LGLHALLRARGVKKSATDHGEKIYLYSKAVRLWHWSN-----ALLFVLLLASGLIN 94
Db
          77 VFSIVSVTAYIALALLSVTISFRI---YKGVIQAIAKSDEGHPFR 118
Qу
              1::| || :| : | : | : | | | | | :|
Db
          95 HFAMVGATAVKSLVAVHEVCGFLLLACWLGFVLINAVGDNGHHYR 139
RESULT 12
T2R8 MOUSE
    T2R8 MOUSE
ΙD
                   STANDARD;
                                  PRT:
                                         246 AA.
    09JKA0;
AC
DТ
    10-OCT-2003 (Rel. 42, Created)
    10-OCT-2003 (Rel. 42, Last sequence update)
DT
     10-OCT-2003 (Rel. 42, Last annotation update)
DT
    Taste receptor type 2 member 8 (T2R8) (Taste receptor family B member
DE
DE
    5) (TRB5) (Fragment).
GN
    TAS2R8.
OS
    Mus musculus (Mouse).
OC
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC
OX
    NCBI TaxID=10090;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=C57BL/6J;
RX
    MEDLINE=20227309; PubMed=10766242;
RA
    Matsunami H., Montmayeur J.-P., Buck L.B.;
    "A family of candidate taste receptors in human and mouse.";
RT
RL
    Nature 404:601-604(2000).
RN
    [2]
RP
    REVIEW.
RX
    MEDLINE=22135574; PubMed=12139982;
    Montmayeur J.-P., Matsunami H.;
RT
     "Receptors for bitter and sweet taste.";
RL
    Curr. Opin. Neurobiol. 12:366-371(2002).
RN
    [3]
RP
    REVIEW.
RX
    MEDLINE=21634924; PubMed=11696554;
RA
    Margolskee R.F.;
RT
     "Molecular mechanisms of bitter and sweet taste transduction.";
```

```
RL
    J. Biol. Chem. 277:1-4(2002).
RN
RP
    REVIEW.
RX
    MEDLINE=22469025; PubMed=12581520;
RA
    Zhang Y., Hoon M.A., Chandrashekar J., Mueller K.L., Cook B., Wu D.,
RA
    Zuker C.S., Ryba N.J.;
RT
    "Coding of sweet, bitter, and umami tastes: different receptor cells
RT
    sharing similar signaling pathways.";
RL
    Cell 112:293-301(2003).
    -!- FUNCTION: Receptor that may play a role in the perception of
CC
        bitterness and is gustducin-linked. May play a role in sensing the
CC
CC
        chemical composition of the gastrointestinal content. The activity
CC
        of this receptor may stimulate alpha gustducin, mediate PLC-beta-2
CC
        activation and lead to the gating of TRPM5.
CC
    -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
    -!- TISSUE SPECIFICITY: Expressed in subsets of taste receptor cells
CC
        of the tongue and palate epithelium and exclusively in gustducin-
CC
        positive cells. Expressed in 15% taste bud cells in circumvallate
        and foliate papillae but only in 2% in fungiform papillae.
CC
CC
    -!- MISCELLANEOUS: Most taste cells may be activated by a limited
CC
        number of bitter compounds; individual taste cells can
CC
        discriminate among bitter stimuli.
CC
    -!- SIMILARITY: Belongs to family T2R of G-protein coupled receptors.
CC
    ______
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
CC
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
    ______
CC
DR
    EMBL; AF247735; AAF64512.1; -.
DR
    MGD; MGI:1890259; Tas2r8.
DR
    InterPro; IPR007960; TAS2R.
DR
    Pfam; PF05296; TAS2R; 1.
KW
    Receptor; G-protein coupled receptor; Transmembrane.
    NON TER
FT
                1
                       1
FT
    DOMAIN
                 1
                       15
                                CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
                16
                       36
                                2 (POTENTIAL).
                37
FT
    DOMAIN
                       59
                               EXTRACELLULAR (POTENTIAL).
FT
    TRANSMEM
               60
                      80
                               3 (POTENTIAL).
FT
    DOMAIN
               81
                    102
                               CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
               103
                    123
                               4 (POTENTIAL).
               124
FT
    DOMAIN
                    153
                               EXTRACELLULAR (POTENTIAL).
FT
    TRANSMEM
               154
                      174
                               5 (POTENTIAL).
                   15.
219
225
FT
    DOMAIN
               175
                               CYTOPLASMIC (POTENTIAL).
FT
    TRANSMEM
               199
                               6 (POTENTIAL).
FT
    DOMAIN
               220
                                EXTRACELLULAR (POTENTIAL).
FT
    TRANSMEM
               226
                     246
                               7 (POTENTIAL).
FT
    NON TER
               246
                      246
              246 AA; 28430 MW; 8B8F96F8A62E4474 CRC64;
SQ
    SEQUENCE
 Query Match
                        10.7%; Score 75.5; DB 1; Length 246;
 Best Local Similarity 27.1%; Pred. No. 3.3;
 Matches 32; Conservative 18; Mismatches 47; Indels 21; Gaps
```

```
17 VYSVSVGMHNLLLLEGRSWQEMDGQKKH---WKDKVVDLLYWRDIKKTGVVFGASLFLLL 73
Qу
                       :|:: || :: | | :: |: |
           40 LYSALMTTRKVLIIFNNSWTVIN----HFNIWLATCLSIFYFLMIAN----FSNSIFLSL 91
Db
           74 SLTVFSIVSVTAYIALALLSV-----TISFRIYKGVIQAIAKSDEG-HPFRAYL 121
QУ
                 1::||| ::| || |
                                     || :||
Db
           92 RWRVKTVVSVTLLMSLLLLFVNVLVINTFIVISVDVYKVNTSYSSHSDNNIHISRIFL 149
RESULT 13
G6PI HELPY
TD
     G6PI HELPY
                   STANDARD;
                                  PRT:
                                         545 AA.
AC
     025781;
     15-JUL-1998 (Rel. 36, Created)
DT
DT
     15-JUL-1998 (Rel. 36, Last sequence update)
DT
     28-FEB-2003 (Rel. 41, Last annotation update)
DE
     Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose
DE
     isomerase) (PGI) (Phosphohexose isomerase) (PHI).
GN
     PGI OR HP1166.
OS
     Helicobacter pylori (Campylobacter pylori).
OC
     Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC
     Helicobacteraceae; Helicobacter.
OX
     NCBI TaxID=210;
RN
     [1]
RP
     SEQUENCE FROM N.A.
RC
     STRAIN=26695 / ATCC 700392;
RX
     MEDLINE=97394467; PubMed=9252185;
     Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA
RA
     Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA
     Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
     Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA
    McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA
RA
    Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
    Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA
    Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA
RA
    Venter J.C.;
RT
    "The complete genome sequence of the gastric pathogen Helicobacter
    pylori.";
RT
RL
    Nature 388:539-547(1997).
CC
    -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC
        phosphate.
CC
    -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC
    -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
    -!- SIMILARITY: Belongs to the GPI family.
CC
CC
    -----
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
CC
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
CC
CC
DR
    EMBL; AE000622; AAD08211.1; -.
DR
    PIR; F64665; F64665.
DR
    HSSP; Q9N1E2; 1HOX.
```

DR

TIGR; HP1166; -.

```
HAMAP; MF 00473; -; 1.
DR
DR
     InterPro; IPR001672; G6P Isomerase.
DR
     Pfam; PF00342; PGI; 1.
DR
     PRINTS; PR00662; G6PISOMERASE.
     PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
DR
     PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
DR
KW
     Isomerase; Gluconeogenesis; Glycolysis; Complete proteome.
FT
     ACT SITE
                 382
                      382
                                  BY SIMILARITY.
FT
     ACT SITE
                                  BY SIMILARITY.
                 510
                        510
SO
     SEQUENCE
                545 AA; 62487 MW; BDC68D1625190236 CRC64;
  Query Match
                          10.7%; Score 75.5; DB 1; Length 545;
  Best Local Similarity
                          25.3%; Pred. No. 7.5;
  Matches
            37; Conservative 20; Mismatches
                                                  44; Indels
                                                                45; Gaps
                                                                             6;
           23 GMHNLL-----LLEGRSWQEMDGQ---KKHWKDKVVDLLYWRDIKKTGVVFGASL 69
Qy
                              ::|:|::|
                                             1
                                                  | | | : | | : |
                                                                   1 11
Db
          411 GHHEILFSNVLAQAQAFMKGKSYEEALGELLFKGLDKDEAKDLAHHR-----VFFGNRP 464
QУ
           70 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD----- 112
                     | :: | :||
                                     | :
                                              : | | | |
          465 SNILLLEKISPSNIGALVALYEHKVFV----QGVIWDINSFDQWGVELGKELAVPILQE 519
Db
Qу
          113 -EGHPFRAYLESEVAISEELVOKYSN 137
               111
                   || :|
                           :: |:: | |
          520 LEGHKSNAYFDSS---TKHLIELYKN 542
Db
RESULT 14
G6PI HELPJ
     G6PI HELPJ
                    STANDARD;
                                   PRT;
                                          545 AA.
     Q9ZK49;
AC
     16-OCT-2001 (Rel. 40, Created)
DT
DT
     16-OCT-2001 (Rel. 40, Last sequence update)
     28-FEB-2003 (Rel. 41, Last annotation update)
DT
DΕ
     Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose
DΕ
     isomerase) (PGI) (Phosphohexose isomerase) (PHI).
GN
     PGI OR JHP1093.
OS
    Helicobacter pylori J99 (Campylobacter pylori J99).
     Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC
OC
    Helicobacteraceae; Helicobacter.
OX
    NCBI TaxID=85963;
RN
    [1]
RP
     SEQUENCE FROM N.A.
RX
    MEDLINE=99120557; PubMed=9923682;
RA
    Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA
     Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA
    Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
    Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA
RA
    Trust T.J.;
     "Genomic sequence comparison of two unrelated isolates of the human
RT
RT
    gastric pathogen Helicobacter pylori.";
    Nature 397:176-180(1999).
RL
CC
    -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC
        phosphate.
CC
    -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC
    -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```
CC
     -!- SIMILARITY: Belongs to the GPI family.
CC
     ______
CC
     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
     between the Swiss Institute of Bioinformatics and the EMBL outstation -
     the European Bioinformatics Institute. There are no restrictions on its
CC
CC
     use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
     or send an email to license@isb-sib.ch).
CC
DR
     EMBL; AE001536; AAD06664.1; -.
DR
     PIR; E71851; E71851.
DR
    HSSP; Q9N1E2; 1HOX.
DR
    HAMAP; MF 00473; -; 1.
DR
    InterPro; IPR001672; G6P Isomerase.
DR
    Pfam; PF00342; PGI; 1.
DR
    PRINTS; PR00662; G6PISOMERASE.
    PROSITE; PS00765; P_GLUCOSE_ISOMERASE 1; 1.
DR
DR
    PROSITE; PS00174; P GLUCOSE ISOMERASE 2; 1.
KW
    Isomerase; Gluconeogenesis; Glycolysis; Complete proteome.
FT
    ACT SITE 382 382 BY SIMILARITY.
    ACT SITE
               510 510
FT
                              BY SIMILARITY.
    SEQUENCE 545 AA; 62302 MW; 7DB544D95FD1D237 CRC64;
SQ
 Query Match 10.6%; Score 74.5; DB 1; Length 545; Best Local Similarity 25.3%; Pred. No. 9.4;
 Matches 37; Conservative 19; Mismatches 45; Indels 45; Gaps
Qу
          23 GMHNLL-----LLEGRSWQEMDGQ---KKHWKDKVVDLLYWRDIKKTGVVFGASL 69
             411 GHHEILFSNVLAQAQAFMKGKSYEEALGELLSKGLDKDEAKDLAHHR-----VFFGNRP 464
Db
Qу
          70 FLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSD------ 112
               :| | | :: | :| | | : | :| | | |
Db
         465 SNILLLEKISPSNIGALVALYEHKVFV----QGVIWDINSFDQWGVELGKELAVPILQE 519
         113 -EGHPFRAYLESEVAISEELVOKYSN 137
Qy
              111 11:1 : 1:: 1 1
Db
         520 LEGHKSNAYFDSS---TRHLIELYKN 542
RESULT 15
YC73 HAEIN
    YC73 HAEIN
                  STANDARD;
                                PRT; 268 AA.
    P44150;
AC
DT
    01-NOV-1995 (Rel. 32, Created)
    01-NOV-1995 (Rel. 32, Last sequence update)
DT
    28-FEB-2003 (Rel. 41, Last annotation update)
DΕ
    Hypothetical protein HI1273.
GN
    HI1273.
OS
    Haemophilus influenzae.
OC
    Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC
    Pasteurellaceae; Haemophilus.
OX
    NCBI TaxID=727;
RN
    [1]
RP
    SEQUENCE FROM N.A.
RC
    STRAIN=Rd / KW20 / ATCC 51907;
```

```
RX
     MEDLINE=95350630; PubMed=7542800;
     Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA
RA
     Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA
     McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA
     Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
     Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA
RA
     Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA
     Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA
     Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
     Venter J.C.;
RA
RT
     "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT
RL
     Science 269:496-512(1995).
     ______
CC
CC
    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
    the European Bioinformatics Institute. There are no restrictions on its
CC
    use by non-profit institutions as long as its content is in no way
CC
    modified and this statement is not removed. Usage by and for commercial
CC
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC
    or send an email to license@isb-sib.ch).
CC
    EMBL; U32807; AAC22921.1; -.
DR
    PIR; F64024; F64024.
DR
DR
    TIGR; HI1273; -.
DR
    InterPro; IPR000051; SAM bind.
KW
    Hypothetical protein; Complete proteome.
SQ
    SEQUENCE 268 AA; 30510 MW; E5B28DA7AADC4D0B CRC64;
 Query Match
                        10.5%; Score 74; DB 1; Length 268;
 Best Local Similarity 25.2%; Pred. No. 5;
 Matches
          29; Conservative 15; Mismatches
                                             49; Indels
                                                            22; Gaps
           8 VSCLRENFAVYSVSVGMHNLLLLEGRSWQEMDGQKKHWKDKVVDLLYWRDIKKTGVVFGA 67
Qy
             : || |:
                          1:1:1
                                             | |: : ||
Db
          94 LDCL----AQFKQKFGLHHLTTFH------KSWADN-----WDDVPQADVVLAS 132
          68 SLFLLLSL-TVFSIVSVTAYIALALLSVTISFRIYKGVIQAIAKSDEGHPFRAYL 121
Qу
                1: | : : | : | | | : | : | : | : | | : | | | | | |
Db
         133 RSTLVDDLDDMIEKLCAKAKKRVFLTSVTQRHFLDEGVFEAIGREDIGFPTYIYL 187
```

Search completed: September 3, 2004, 16:06:16 Job time: 6.8747 secs